

В.Ф. Гантмахер

**ЭЛЕКТРОНЫ
В НЕУПОРЯДОЧЕННЫХ
СРЕДАХ**



В.Ф. Гантмахер

ЭЛЕКТРОНЫ В НЕУПОРЯДОЧЕННЫХ СРЕДАХ

*Издание третье,
исправленное и дополненное*



МОСКВА[®]
ФИЗМАТЛИТ
2013

УДК 538.9
ББК 22.37
Г 19

Гантмахер В.Ф. **Электроны в неупорядоченных средах.** — М.: ФИЗМАТЛИТ, 2013. — 288 с. — ISBN 978-5-9221-1487-5.

Книга предназначена студентам старших курсов и аспирантам, специализирующимся в области физики твердого тела, а также научным сотрудникам и всем, кто профессионально нуждается в понимании основ физических процессов, управляющих поведением электронов в твердых телах. Она написана с минимумом математики. Основное внимание уделено обсуждению физической сущности явлений и выявлению глубинных связей и аналогий. В настоящее 3-е издание добавлены главы о квантовом эффекте Холла и о квантовых фазовых переходах. Некоторые главы переработаны. Ко всем главам добавлены вопросы и задачи.

Научное издание

ГАНТМАХЕР Всеволод Феликсович

ЭЛЕКТРОНЫ В НЕУПОРЯДОЧЕННЫХ СРЕДАХ

Редактор *Е.С. Артоболевская*

Оригинал-макет: *Д.В. Горбачев*

Оформление переплета: *Н.Л. Лисицына*

Подписано в печать 27.06.2013. Формат 60×90/16. Бумага офсетная. Печать офсетная.
Усл. печ. л. 18. Уч.-изд. л. 19.8. Тираж 300 экз. Заказ №

Издательская фирма «Физико-математическая литература»

МАИК «Наука/Интерпериодика»

117997, Москва, ул. Профсоюзная, 90

Е-mail: fizmat@maik.ru, fmlsale@maik.ru;

<http://www.fml.ru>

Отпечатано с электронных носителей
в ООО «Чебоксарская типография № 1»
428019, г. Чебоксары, пр. И. Яковлева, 15
Тел.: (8352) 28-77-98, 57-01-87
Сайт: www.volga-print.ru

ISBN 978-5-9221-1487-5



9 785922 114875

ISBN 978-5-9221-1487-5

© ФИЗМАТЛИТ, 2013

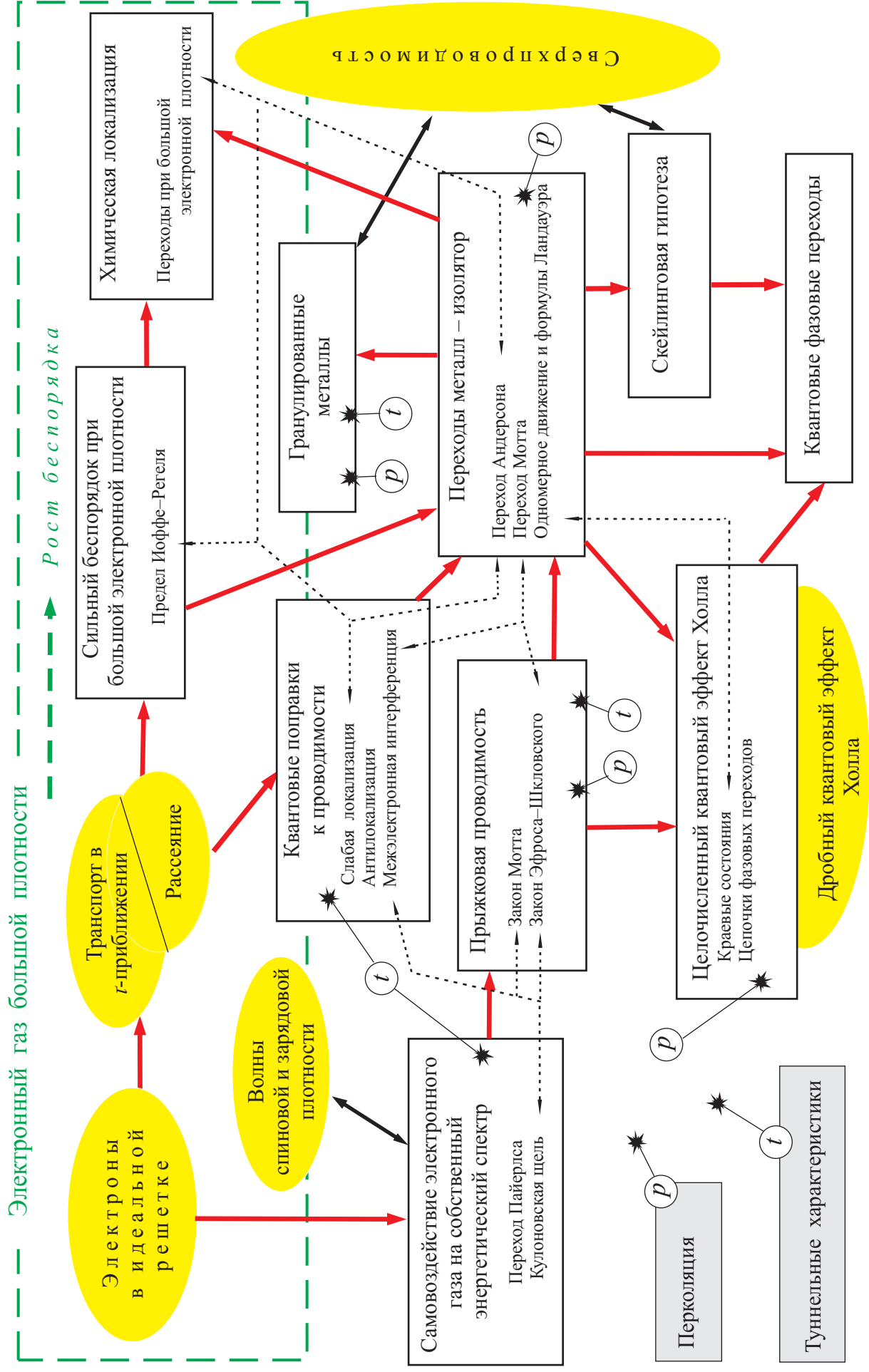
© В. Ф. Гантмахер, 2013

ОГЛАВЛЕНИЕ

Предисловие	6
Литература по разделам физики металлов, не вошедшим в книгу	9
Благодарности	10
Глава 1. Металлы с сильным беспорядком	11
1.1. Дифракционная теория электронного транспорта в жидких металлах	12
1.2. Правило Мооя	17
1.3. Насыщение сопротивления	19
1.4. Предел Иоффе–Регеля при большой электронной плотности	23
1.5. Контрольные вопросы и задачи	25
Глава 2. Квантовые поправки к проводимости	27
2.1. Слабая локализация	28
2.2. Влияние магнитного поля на слабую локализацию	35
2.3. Антилокализация	41
2.4. Межэлектронная интерференция	48
2.5. Сравнительный анализ квантовых поправок	53
2.6. Контрольные вопросы и задачи	55
Глава 3. Влияние межэлектронного взаимодействия на электронный энергетический спектр	58
3.1. Переход Пайерлса	58
3.2. Структура примесной зоны при слабом легировании	60
3.3. Кулоновская щель	67
3.4. Контрольные вопросы и задачи	72
Глава 4. Прыжковая проводимость	74
4.1. Локализованные состояния и переходы между ними	74
4.2. Прыжки на ближайшие центры	77
4.3. Прыжки с переменной длиной прыжка	80
4.4. Экспериментальные наблюдения прыжковой проводимости	83
4.5. Контрольные вопросы и задачи	89

Глава 5. Переходы металл–изолятор	90
5.1. Переход Андерсона	92
5.2. Модель структурного беспорядка	94
5.3. Переход Мотта	97
5.4. Минимальная металлическая проводимость?	102
5.5. Формула Ландауэра для одномерных (1D) систем	104
5.6. Локализация и роль корреляций в 1D-системах	109
5.7. Микроволновое моделирование	116
5.8. Контрольные вопросы и задачи	119
 Глава 6. Скейлинговая гипотеза	122
6.1. Обоснование и формулировка скейлинговой гипотезы	122
6.2. Трехмерные (3D) системы	128
6.3. Критическая окрестность перехода	130
6.4. Мультифрактальные волновые функции	136
6.5. Двумерные (2D) системы	141
6.6. Квазиодномерные (q1D) системы	146
6.7. Скейлинг и взаимодействия	150
6.8. Контрольные вопросы и задачи	154
 Глава 7. Химическая локализация	156
7.1. Интерметаллические комплексы в двухкомпонентных расплавах	157
7.2. Квазикристаллы	163
7.3. Модели электронной структуры в системах с металлическими ква- зимолекулами-ловушками	169
7.4. Контрольные вопросы и задачи	172
 Глава 8. Гранулированные металлы	174
8.1. Морфология и классификация	174
8.2. Кулоновская блокада и переход металл–изолятор	180
8.3. Фрактально-гранулированные металлы	187
8.4. Контрольные вопросы и задачи	192
 Глава 9. Целочисленный квантовый эффект Холла	193
9.1. Спектр и динамика двумерных электронов в сильном магнитном поле	194
9.2. Экспериментальные наблюдения целочисленного КЭХ	202
9.3. Механизм образования плато	206
9.4. Краевые каналы	213

9.5. Плотность состояний электронного 2D-газа в магнитном поле	218
9.6. Цепочки фазовых переходов	221
9.7. Двухпараметрический скейлинг	227
9.8. Контрольные вопросы и задачи	234
 Глава 10. Квантовые фазовые переходы	 236
10.1. Параллели и различия между классическим и квантовым фазовыми переходами	237
10.2. Критическая окрестность квантового перехода	244
10.3. Квантовые переходы металл–изолятор	248
10.4. Квантовые переходы между разными состояниями холловской жидкости	252
 Приложение А. Элементы теории перколяции	 257
А.1. Аппроксимация эффективной среды	257
А.2. Перколяционные пороги	260
А.3. Окрестность перколяционного перехода	264
А.4. Пример: электропроводность сильно неоднородной среды	267
А.5. Контрольные вопросы и задачи	270
Приложение Б. Туннельные характеристики	271
Б.1. Сканирующая туннельная спектроскопия	281
Б.2. Контрольные вопросы и задачи	284
 Указатель материалов	 286
Предметный указатель	287



Электронны в неупорядоченных средах (Путеводитель)

ПРЕДИСЛОВИЕ

Книга, которую вы держите в руках, это скорее не учебник, а путеводитель. Как известно, путеводитель не может заменить путешествие. Но он подсказывает, куда следует повернуть, где задержаться, на что обратить внимание, над чем задуматься. Когда вы находитесь в стране с незнакомым языком, путеводитель должен быть еще и разговорником, подсказывая, как спросить дорогу или как понять надпись на указателе.

Путеводитель должен начинаться с обзорной карты. Роль карты для страны «Электроны в неупорядоченных средах» выполняет схема-оглавление, разъясняющая логику «административного деления», специализацию и внутренние связи между главами-«провинциями». В настоящем, третьем издании, их двенадцать. (В первом издании было десять — здесь добавлены главы о квантовом эффекте Холла и о квантовых фазовых переходах; кроме того, подверглись переработке гл. 5 и 6, в меньшей степени гл. 2, ко всем главам добавлены вопросы и задачи. Остальные изменения незначительны.) Помимо десяти естественных исторически сложившихся «провинций», которые обозначены белыми прямоугольниками, на карте имеются еще две, закрашенные серым, появившиеся в соответствии с «методическим принципом» и вынесенные в приложение. Раздел «Элементы теории перколяции» представляет собой краткое изложение основных понятий этой математической дисциплины. Появился здесь он потому, что эта молодая теория, хотя и широко используется, до сих пор не включена в университетские курсы математики. В разделе «Туннельные характеристики» описаны эксперименты, имеющие непосредственное отношение одновременно к самым различным обсуждаемым явлениям: межэлектронной интерференции, кулоновской щели, прыжковой проводимости, переходам металл–изолятор. Выделение этого материала в отдельный раздел позволяет избежать повторений. Предполагается, что читатель, единожды разобравшись с сутью эксперимента, будет возвращаться в этот раздел из разных мест, чтобы посмотреть на экспериментальные кривые. Специальные значки, смысл которых нетрудно понять, показывают, в каких разделах используются сведения, собранные в обоих приложениях.

На обзорную карту обычно наносят и сопредельные территории. У нас они окрашены в желтый цвет. Указаны и средства сообщения с ними: по названию в конце Предисловия можно найти соответствующие ссылки на учебную и обзорную литературу. Предполагается, что

читатель знает и понимает то, что подразумевается под «Электронами в идеальной решетке», «Транспортом в τ -приближении» и «Рассеянием» или в крайнем случае может воспользоваться соответствующими учебниками. Знакомство со сверхпроводимостью или с волнами зарядовой плотности, вообще говоря, не обязательно.

Карта сама несет информативную нагрузку, являясь чем-то большим, чем просто оглавлением. При этом следует помнить, что, поскольку страна виртуальная, то карта субъективна: подобно географическим картам древности, ее вид зависит от вкусов и взглядов составителя.

На предлагаемой карте не указана большая сопредельная территория, которая называется «Взаимодействующие электроны». Это страна будущего. Многие открытия в ней, ее освоение, прокладывание дорог и застройка еще только предстоят. По существу «Сверхпроводимость» с ее куперовскими парами и эффектом Джозефсона и «Двумерный электронный газ» с дробным квантовым эффектом Холла и композитными фермионами — это окраинные провинции этой страны, освоенные одна более, другая менее. Возможно, что в названии этой книги следовало бы написать «невзаимодействующие электроны...». Но это было бы тоже неточно, потому что кулоновское межэлектронное взаимодействие в ней обсуждается довольно активно. Оно является существенным для кулоновской щели, переходов Пайерлса и Мотта и других разделов, которые можно легко определить по пунктирным стрелкам на карте. Но проблемы межэлектронного взаимодействия не исчерпываются законом Кулона. Эту книгу можно рассматривать как плацдарм для их изучения.

Однако, надо иметь в виду, что и плацдарм еще освоен не до конца. Вам предстоит путешествие по сравнительно новой и развивающейся стране. В одних местах строительство в основном завершилось, и пейзаж вряд ли изменится в будущем. В других можно оказаться на стройплощадке или на территории, подлежащей реконструкции, так что через несколько лет многое будет выглядеть не так, как сейчас.

Особенности предмета «Электроны в неупорядоченных средах» в том, что нет ни одного большого учебника, в котором он был бы целиком описан. Поэтому по ходу изложения мы указываем книги, обзоры, а иногда даже оригинальные статьи, где соответствующий материал изложен наиболее просто и понятно (а этот выбор опять-таки весьма субъективен). Такая система ссылок ни в какой мере не отражает приоритетов в получении новых результатов. Например, любой график с экспериментальными результатами сопровождается ссылкой на оригинальную работу, чтобы можно было посмотреть детали эксперимента. Но сами результаты отбирались по принципу «зрелищности» и могут быть не первыми по времени. Теоретические работы, содержащие новые оригинальные результаты, часто мало пригодны для первого ознакомления с предметом из-за сложного

математического аппарата. Такие работы здесь тоже не упоминаются. Но встречаются ссылки на работы, содержащие простые наглядные модели, не получившие дальнейшего развития и не вошедшие в обзоры.

Засим, счастливого пути! Рекомендуем двигаться по жирным красным стрелкам, придерживаясь преимущественного направления слева направо. Двухсторонние пунктирные стрелки указывают на внутренние взаимосвязи, а сплошные — на связи с внешним миром. Одного канонического маршрута нет. Каждому — по потребностям.

Литература по разделам физики металлов, не вошедшим в книгу

Электроны в идеальной решетке

1. *Абрикосов А. А.* Основы теории металлов. — М.: Физматлит, 2006.
2. *Ашкрофт Н., Мермин Н.* Физика твердого тела. — М.: Мир, 1979 (перевод книги *Ashcroft N. W., Mermin N. D.* Solid State Physics. — Holt, Rinehart and Winston, 1969).
3. Физика металлов. 1. Электроны (ред. Дж. Займан). — М.: Мир, 1972 (перевод книги *The physics of metals. 1. Electrons* (ed. J. Ziman). — Cambridge Univ. Press, 1969).

Транспорт в τ -приближении и рассеяние

4. *Абрикосов А. А.* Основы теории металлов. — М.: Физматлит, 2006.
5. *Гантмахер В. Ф., Левинсон И. Б.* Рассеяние носителей тока в металлах и полупроводниках. — М.: Наука, 1984.

Волны спиновой и зарядовой плотности

6. *Уайт Р., Джебелл Т.* Дальний порядок в твердых телах. — М.: Мир, 1982 (перевод книги *White R. M., Geballe T. H.* Long range order in solids. — Academic Press, 1979).
7. *Grüner G.* Density Waves. — Perseus Books, 2000.
8. *Grüner G.* // Rev. Mod. Phys. **60**, 1129 (1988); **66**, 1 (1994).

Сверхпроводимость

9. *Абрикосов А. А.* Основы теории металлов. — М.: Физматлит, 2006.
10. *Шмидт В. В.* Введение в физику сверхпроводников. — М.: Наука, 1982; 2-е изд. — М.: МЦНМО, 2000.

Двумерный электронный газ

11. *Демиховский В. Я., Вугальтер Г. А.* Физика квантовых низкоразмерных структур. — М.: Логос, 2000.
12. Physics of low dimensional structures (ed. B. Butcher, N.H. March, M.P. Tosi). — Plenum Press, 1993.

БЛАГОДАРНОСТИ

Эта книга возникла из курса, который автор читал в течении ряда лет студентам Московского физико-технического института и Московского государственного университета. Все слушатели курса своими вопросами на лекциях и ответами на экзаменах способствовали его оптимизации. Автор благодарен Ю. Гальперину, В. Долгополову, Э. Рашбе и С. Студеникину, прочитавшим рукопись и сделавшим много ценных замечаний. Особая благодарность Д. Хмельницкому, дискуссии с которым привели к уточнению многих утверждений и даже к появлению новых параграфов. При написании главы про квантовый эффект Холла автор пользовался советами и разъяснениями В. Долгополова и С. Мурзина.

Глава 1

МЕТАЛЛЫ С СИЛЬНЫМ БЕСПОРЯДКОМ

Основным фактором, определяющим законы движения делокализованных электронов в кристаллах, является дальний порядок. Именно он обеспечивает в кристалле интерференцию после рассеяния электронов от отдельных атомов, которая полностью гасит рассеяние и допускает стационарное распространение электронных волн почти со всеми волновыми векторами $\mathbf{k}$. Исключения составляют лишь волны, у которых волновой вектор удовлетворяет условию

$$k^2 = (\mathbf{k} - \mathbf{K}_m)^2, \quad (1.1)$$

где $\mathbf{K}_m$ — произвольный вектор обратной решетки. Для электронов с такими волновыми векторами рассеяние является резонансным, и такие электронные волны вообще не могут распространяться в кристалле. Все же остальные электроны рассеиваются лишь на отклонениях от периодичности, что учитывается понятием длины свободного пробега l . В таком контексте l имеет смысл расстояния, которое электрон проходит между двумя последовательными независимыми актами рассеяния. Естественно, что рассеяние должно быть не слишком частым, чтобы длина l была больше электронной длины волны $\lambda \approx 2\pi/k_F$:

$$k_F l \gtrsim 1. \quad (1.2)$$

Под k_F здесь подразумевается радиус ферми-сферы

$$k_F = (3\pi^2 n)^{1/3}, \quad (1.3)$$

определяемый через концентрацию свободных носителей n . Ограничение снизу на длину пробега $l \gtrsim 1/k_F$ называют пределом Иоффе–Регеля. Обоснование этого неравенства, заимствованное из классической физики, звучит так: отрезок синусоиды размером меньше длины волны уже нельзя считать синусоидой. Квантовомеханический аргумент, по сути дела аналогичный, исходит из соотношения неопределенности $\Delta k \Delta x \sim 1$. Поскольку Δk заведомо меньше k_F , минимально возможная неопределенность траектории электрона $\Delta x \gtrsim 1/k_F$. Естественно, что расстояние между двумя актами рассеяния вдоль этой траектории должно быть больше этой неопределенности.

В этой главе мы будем полагать, что имеем дело с «настоящими» металлами, у которых концентрация n такова, что среднее расстояние между носителями порядка среднего расстояния a между атомами:

$$n^{-1/3} \lesssim (n^*)^{-1/3} = a \approx 3 \text{ \AA}, \quad n \gtrsim 4 \cdot 10^{22} \text{ см}^{-3}. \quad (1.4)$$

Такой металл мы будем называть стандартным. В веществах со столь большой электронной плотностью при наличии дальнего порядка ферми-поверхности обычно имеют довольно сложную форму. Однако для оценок всегда можно пользоваться моделью ферми-сферы (1.3).

Длину пробега l можно оценить из величины удельной проводимости σ . Подставив ферми-радиус (1.3) в выражение для σ , получим формулу Друде

$$\sigma = \frac{ne^2 l}{\hbar k_F} = (3\pi^2)^{-2/3} \frac{e^2}{\hbar} n^{1/3} (k_F l). \quad (1.5)$$

Согласно неравенству Иоффе–Регеля (1.2), минимальное значение безразмерного параметра $k_F l$ равно единице. Тогда из (1.5) следует, что удельное сопротивление не может быть больше, чем $\rho^* \approx 10(\hbar/e^2)a$. Конечно, получающееся из соотношения (1.5) численное значение $\rho^* \approx 1000$ мкОм·см весьма приблизительно и его следует уточнить экспериментально. Как мы увидим (см. ниже, рис. 1.3), из экспериментов следует значение ρ^* для стандартного металла

$$\rho^* \approx (200 \div 300) \text{ мкОм} \cdot \text{см}, \quad (1.6)$$

которым мы и будем пользоваться.

Само существование предельного значения ρ^* позволяет поставить два вопроса.

1. Каковы транспортные свойства сильно разупорядоченного стандартного металла при большом ρ , которое тем не менее $\rho < \rho^*$?

2. Нельзя ли преодолеть предел ρ^* и изготовить стандартный металл с $\rho > \rho^*$? В простейшей модели (1.5) это бы означало, что $l < a$, т.е. что электрон заперт и движется в пределах одной элементарной ячейки или вблизи одного атома и что металл превратился в изолятор. Переход металл–изолятор под влиянием беспорядка известен (переход Андерсона), но он всегда наблюдается в системах с электронной плотностью, существенно меньшей, чем (1.4). Поэтому второй вопрос можно переформулировать так: возможен ли переход Андерсона в системах с большой электронной плотностью (1.4)?

В этой главе мы обсудим первый вопрос и начнем обсуждение второго, а завершим это обсуждение в гл. 7.

1.1. Дифракционная теория электронного транспорта в жидких металлах ¹⁾

Когда статические дефекты периодической решетки расположены далеко друг от друга, рассеяние на каждом из них происходит независимо. При постепенном увеличении концентрации дефектов их потенциалы должны начать перекрываться. Тогда в пространстве уже

¹⁾ Материал этого параграфа обсуждается также в книгах [1] и [2].

нельзя выделить области, свободные от рассеивающих полей, и четко указать, где какой статический дефект ответственен за рассеяние. Увеличение беспорядка требует разработки нового подхода для описания его последствий. Такой подход предложен в теории жидких металлов Займана.

В жидкости сохраняется только ближний порядок: ближайшее окружение каждого атома расположено почти так же, как в кристалле. Но именно из-за этого «почти» при переходе к атомам, более удаленным от исходного, неопределенность их расположения относительно исходного атома нарастает, так что дальний порядок отсутствует. Кажется бы при отсутствии дальнего порядка, когда каждый атом рассеивает независимо, условие (1.2) должно нарушиться. Однако довольно часто, в частности в моноэлементных жидких металлах, это не так. Об этом свидетельствует величина удельного сопротивления $\rho = 1/\sigma$. В формуле (1.5) выразим концентрацию носителей n через валентность Z , т.е. число свободных электронов на атом, и концентрацию атомов $N = 1/a^3$, определяемую из удельного веса расплава. Благодаря этим соотношениям, из величины удельной проводимости непосредственно определяется отношение l/a . В большинстве моноэлементных жидких металлов это отношение больше 5, а в щелочных металлах даже больше 100 (кроме Li, у которого $l/a \approx 13$). Это означает, что сечение рассеяния на отдельных атомах не очень большое, в несколько раз меньше, чем a^2 . Причина ослабленного рассеяния — в большой электронной плотности и, как следствие, в сильном экранировании. Каждый электрон чувствует не истинный потенциал иона, а лишь его перенормированную малую часть, сохранившуюся после экранирования остальными электронами. Этот остаток называется псевдопотенциалом, и в дальнейшем в этом параграфе речь будет идти только о нем, хотя приставку «псевдо» мы будем для краткости опускать.

Излагаемая теория описывает рассеяние электронов на слабом, но протяженном хаотическом потенциале. Специфика задачи в том, что рассеивателем является потенциал $V(\mathbf{r})$, охватывающий весь объем. При этом предполагается, что энергетический спектр изотропен и энергия электрона зависит только от модуля его волнового вектора: $\varepsilon = \varepsilon(k)$, а волновые функции имеют вид немодулированных плоских волн $\exp(i\mathbf{k}\mathbf{r})$. При таких волновых функциях матричный элемент перехода

$$\int \psi_2^* V(\mathbf{r}) \psi_1 d\mathbf{r} = \int e^{i(\mathbf{k}_1 - \mathbf{k}_2)\mathbf{r}} V(\mathbf{r}) d\mathbf{r} = V(\mathbf{q}) \quad (1.7)$$

есть фурье-компонента рассеивающего потенциала с аргументом, равным изменению волнового вектора при рассеянии: $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_2$. Потенциал $V(\mathbf{r})$ складывается из потенциалов отдельных атомов $v(\mathbf{r} - \mathbf{R}_i)$, расположенных в точках $\mathbf{R}_i$:

$$V(\mathbf{r}) = \sum_{\mathbf{R}_i} v(\mathbf{r} - \mathbf{R}_i). \quad (1.8)$$

Соответственно, фурье-компонента потенциала $V(\mathbf{r})$

$$V(\mathbf{q}) = \sum_{\mathbf{R}_i} \int e^{i\mathbf{q}\mathbf{r}} v(\mathbf{r} - \mathbf{R}_i) d\mathbf{r} = \sum_{\mathbf{R}_i} e^{i\mathbf{q}\mathbf{R}_i} \int e^{i\mathbf{q}\mathbf{r}} v(\mathbf{r}) d\mathbf{r} = v(\mathbf{q}) \sum_{\mathbf{R}_i} e^{i\mathbf{q}\mathbf{R}_i} \quad (1.9)$$

выражается через фурье-компоненту потенциала отдельного атома $v(\mathbf{q})$.

Заметьте. Форма записи (1.8) потенциала $V(\mathbf{r})$ предполагает, что волновые функции различных ионов перекрываются не очень сильно, т. е. что нет пар с очень малыми расстояниями $\mathbf{R}_i - \mathbf{R}_j$.

Сравните. Формально точно такой же потенциал фигурирует и в модели структурного беспорядка, обсуждаемой в гл. 5 в связи с переходами Андерсона. Но там у потенциала нет приставки «псевдо», ямы $v(\mathbf{r} - \mathbf{R}_i)$ глубокие, и электроны могут в принципе сидеть каждый в своей яме.

Поскольку вероятность рассеяния выражается через квадрат матричного элемента, нужно вычислить квадрат $V^2(\mathbf{q})$. Пусть объем, занимаемый жидким металлом, равен единице, а концентрация ионов N . Тогда

$$|V(\mathbf{q})|^2 = |v(\mathbf{q})|^2 \sum_{\mathbf{R}_i, \mathbf{R}_j} e^{i\mathbf{q}(\mathbf{R}_i - \mathbf{R}_j)} = |v(\mathbf{q})|^2 \left(N + \sum_{\mathbf{R}_i, \mathbf{R}_j, i \neq j} e^{i\mathbf{q}(\mathbf{R}_i - \mathbf{R}_j)} \right). \quad (1.10)$$

Последнее преобразование справедливо благодаря тому, что все N диагональных элементов $i = j$ двойной суммы равны 1. Зафиксируем некоторый ион $j = j_0$, перенесем начало координат в точку $\mathbf{R}_{j_0}$ и усредним сумму

$$\sum_{\mathbf{R}_i} e^{i\mathbf{q}(\mathbf{R}_i - \mathbf{R}_{j_0})} \equiv \sum_{\mathbf{R}_i} e^{i\mathbf{q}\mathbf{R}'_i}, \quad \mathbf{R}'_i = \mathbf{R}_i - \mathbf{R}_{j_0},$$

по всем возможным конфигурациям $\{\mathbf{R}_i\}$. Тогда суммирование по $\mathbf{R}_j$ можно заменить умножением на N . Обозначив результат усреднения чертой сверху, получим

$$\overline{|V(\mathbf{q})|^2} = |v(\mathbf{q})|^2 N \left(1 + \overline{\sum_{\mathbf{R}_i} e^{i\mathbf{q}\mathbf{R}_i}} \right) = |v(\mathbf{q})|^2 N S(q). \quad (1.11)$$

Выражение в скобках, обозначенное $S(q)$, называется структурным фактором. Для преобразования $S(q)$ и выяснения его физического смысла введем вероятность $NP(r) d^3r$ нахождения иона в объеме d^3r при условии, что другой ион находится в начале координат $r = 0$; интеграл по единичному объему $\int P(r) d^3r = 1$. Если положения всех ионов статистически независимы, то $P(r) \equiv 1$. Наличие спадающих с r корреляций ближнего порядка отражается на значениях P при значениях r порядка a , а при $r \rightarrow \infty$ всегда функция $P(r) \rightarrow 1$. Из-за

этого фурье-образ $P(q)$ функции $P(r)$ содержит сингулярность в виде дельта-функции: $P(q) = \delta(q) + P'(q)$, где $P'(q)$ — функция, регулярная в точке $q = 0$. Но родственная $P(r)$ парная корреляционная функция $Q(r) = P(r) - 1$, которая на больших r стремится к нулю, не имеет этого недостатка.

Заменяя в определении (1.11) структурного фактора $S(q)$ сумму, усредненную по всем позициям атомов $\mathbf{R}_i$, на интеграл по d^3r , в котором вероятность $P(r)$ является весовым множителем. С точностью до дельта-функции, на которую различаются фурье-образы функций $P(r)$ и $Q(r)$, имеем

$$\begin{aligned} S(q) &= 1 + \overline{\sum_{\mathbf{R}_i} e^{i\mathbf{q}\mathbf{R}_i}} = 1 + N \int e^{i\mathbf{q}\mathbf{r}} Q(r) d^3r = \\ &= 1 + 2\pi N \int \int e^{iqr \cos \theta} \sin \theta d\theta Q(r) r^2 dr = \\ &= 1 + 4\pi N \int_0^\infty Q(r) \frac{\sin qr}{qr} r^2 dr. \quad (1.12) \end{aligned}$$

Функция $S(q)$ появляется во всех дифракционных задачах, в частности, в задачах о рассеянии нейтронов и рентгеновских лучей, и может быть извлечена из соответствующих экспериментов. Как и $P(r)$, она содержит всю информацию о корреляциях в положениях ионов. Появление здесь этой функции подчеркивает, что весь подход базируется на предположении, что волновые функции электронов это плоские волны. Поэтому эту теорию часто называют спектральной или дифракционной теорией транспорта в жидких металлах.

Обычный вид функции $S(q)$ показан на рис. 1.1: после нескольких затухающих осцилляций она выходит на асимптоту $S = 1$. Масштаб

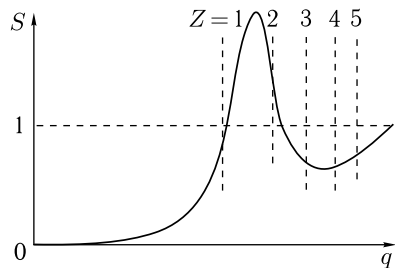


Рис. 1.1. Схематический график функции $S(q)$, на котором отмечены величины $2k_F$ при разном количестве Z свободных электронов на атом

этих осцилляций по оси ординат зависит от корреляций: чем они слабее, тем амплитуда осцилляций меньше. Для системы со статистически независимыми положениями ионов $S(q) = 1$. Масштаб по оси абсцисс задается средним расстоянием между ионами $a \approx N^{-1/3}$. Значение аргумента в первом максимуме определяется радиусом a_1 первой координационной сферы $1/q_1 \approx a_1 \approx a$. Аналогично $1/q_2 \approx a_2 \approx 2a$ и т.д. Благодаря этому, из соотношений (1.5) следует, что на оси абсцисс можно указать точки $q = 2k_F$ для различных Z (вертикальные пунктирные линии на рис. 1.1) и их расположение относительно максимумов

функции $S(q)$. Именно это позволило сделать из теории нетривиальные выводы, поддающиеся экспериментальной проверке.

Как известно, входящая в формулу (1.5) величина l аккумулирует результаты всех актов упругого рассеяния $|\mathbf{k}_1| = |\mathbf{k}_2| = k_F$ на разные углы $\theta = 2 \arcsin(q/2k_F)$ (см. рис. 1.2) в соответствии с формулой

$$\frac{1}{l} \propto \int_0^\pi \overline{|V(q)|^2} (1 - \cos \theta) \sin \theta d\theta \propto \int_0^1 S(2k_F x) v^2(2k_F x) x^3 dx, \quad x = \frac{q}{2k_F}. \quad (1.13)$$

Из-за множителя x^3 в подынтегральном выражении в (1.13) основной вклад в интеграл вносит область $x \approx 1$, т. е. рассеяние «почти назад» на углы $\theta \approx \pi$. Это означает, что при вычислении величины $1/l$ существенна не вся функция $S(q)$, а лишь ее значения в окрестности точки $q = 2k_F$ (см. рис. 1.1). Отсюда следует чрезвычайно сильное и элегантное утверждение.

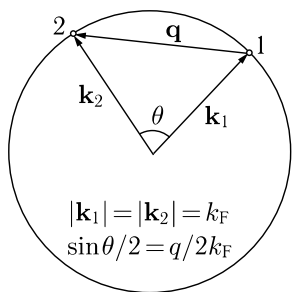


Рис. 1.2. К формуле (1.13): связь между углом упругого рассеяния θ и модулем переданного импульса q

Очевидно, что при повышении температуры корреляции ослабевают и система ионов становится все более хаотичной. Поэтому значения функции $S(q)$ при каждом фиксированном q с ростом температуры приближаются к единице. Как видно из рис. 1.1, в металлах с $Z = 1$ (жидкие Na, K, Rb) и $Z = 3$ (Al, Ga, In) величина $S(2k_F)$ при этом будет расти, а в металлах второй группы с $Z = 2$ (жидкие Zn, Cd, Hg) —

уменьшаться. Соответственно по-разному должны меняться с температурой величина $1/l$ и пропорциональное ей удельное сопротивление ρ . Таким образом, дифракционная теория предсказывает нетривиальную зависимость от валентности знака температурного коэффициента сопротивления жидких металлов: у щелочных и трехвалентных металлов сопротивление должно расти с ростом температуры, а у двухвалентных — падать. Как видно из табл. 1.1, где приведены данные о температурном коэффициенте сопротивления для ряда металлов, дело именно так и обстоит. (Все металлы, фигурирующие в таблице, имеют сравнительно низкие температуры плавления: ниже 200°C у всех, кроме Zn (420°C) и Cd (320°C).)

Почувствителен еще один аспект результатов, получающихся в дифракционной теории. Оказалось, что увеличение беспорядка может приводить к уменьшению сопротивления. Необходимой предпосылкой для этого является перекрытие отдельных рассеивателей, эффект от которых был бы аддитивен, будь они изолированы. Кроме того существенно, чтобы тепловые фононы добавлялись к фону сильного статического беспорядка, который уже разрушил дальний порядок и анизотропию.

Таблица 1.1

Металл	Валентность	ρ , мкОм · см	$\frac{d \ln \rho}{d \ln T}$
Li	1	25	0,6
Na	1	10	0,85
K	1	13	0,76
Rb	1	22	0,70
Cs	1	37	0,69
Zn	2	37	−0,24
Cd	2	34	−0,22
Hg	2	91	−0,10
Ga	3	26	0,14
In	3	33	0,16

Однако весьма примечателен сам факт наличия такого эффекта, получаемого в рамках приближения кинетического уравнения.

Заметьте. Рост сопротивления с понижением температуры происходит также в условиях слабой локализации (см. гл. 2). Но там он обусловлен квантовыми поправками к проводимости, которые появляются за пределами приближения кинетического уравнения.

1.2. Правило Моойа

Успех дифракционной теории показывает, что можно по-прежнему опираться и ссылаться на величину параметра l даже тогда, когда l уже трудно назвать расстоянием, которое электрон проходит между двумя последовательными точечными актами рассеяния. Из хода вычислений скорее следует, что рассеяние при сильном беспорядке представляет собой как бы непрерывный процесс, количественно оцениваемый длиной l . В этих условиях решающее значение приобретает значение безразмерного параметра $l/a \approx k_F l$.

Блестящее подтверждение выводов дифракционной теории в моноэлементных металлических расплавах привело к тому, что эту теорию стали широко применять при обработке данных по электросопротивлению разупорядоченных высокоомных металлических сплавов по следующей схеме: по знаку температурного коэффициента сопротивления определяется среднее количество носителей на атом, что далее используется как аргумент при определении электронного спектра, перекрытия зон, положений ферми-уровня и т. д. Оказалось, однако, что во многих случаях простейшие оценки по формулам (1.5) показывают, что $l \lesssim a$. Тогда плоские волны являются плохой аппроксимацией волновых функций и применение дифракционной теории становится неправомерным.

В этих условиях следует обратиться в первую очередь к эксперименту. На рис. 1.3 собраны данные о температурном коэффициенте сопротивления $\bar{\alpha} = R^{-1}(dR/dT)$ высокорезистивных металлических сплавов для более ста различных материалов (величина $\bar{\alpha}$ имеет размерность обратной энергии). На основании этих данных был сделан вывод о существовании корреляции между величиной удельного со-

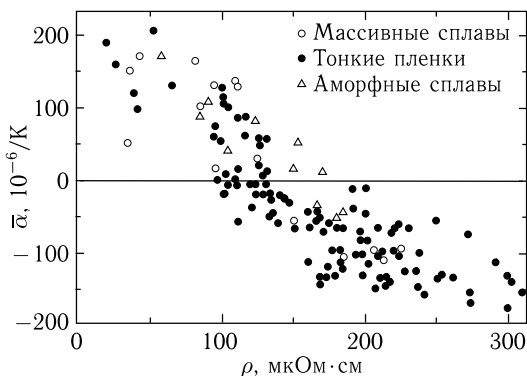


Рис. 1.3. Правило Мооя: корреляция между удельным сопротивлением и его температурным коэффициентом в высокоомных сплавах [3]

противления и температурным коэффициенте сопротивления и сформулировано эмпирическое правило Мооя: *Металлические материалы с удельным сопротивлением, меньшим 100–150 мкОм·см, имеют преимущественно положительный температурный коэффициент сопротивления, а с большим — отрицательный.* Другими словами, если статический беспорядок создает «слишком» большое удельное сопротивление, то температура его немного уменьшает.

Заметьте. Речь идет лишь о материалах, состоящих только из металлических атомов, валентные электроны которых слабо связаны с ионными остовами и поэтому делокализованы, т. е. лишь о стандартных металлах в смысле (1.4). Сюда заведомо не относятся окислы металлов, полуметаллы типа Bi, соединения типа MoGe, высокотемпературные сверхпроводники и т. п.

С другой стороны. Как видно из табл. 1.1, почти все моноэлементные жидкие металлы, являвшиеся объектами приложения дифракционной теории, имеют значения $\rho \lesssim 50$ мкОм·см, так что эти материалы тоже не подпадают под действие правила Мооя.

На графике рис. 1.3 не случайно отсутствуют данные о сплавах с $\rho > 300$ мкОм·см. Таких сплавов практически нет. Сотни исследованных композиций металлических сплавов, десятки различных методов изготовления разупорядоченных систем — все приводят к материалам с удельным сопротивлением, меньшим 300 мкОм·см. Хотя эта

величина в 3–4 раза меньше, чем максимальное сопротивление ρ^* , полученное из оценки на основании формул (1.5), по всей видимости именно его следует считать максимально возможным сопротивлением стандартного металла, которое получается в пределе $l \approx a$.

Наличие верхнего предела удельного сопротивления стандартного металла, как и правило Моойа, по-видимому обусловлены глубокими физическими причинами, связанными с экранированием в системе с большой плотностью свободных носителей. Вероятно, уменьшая амплитуду случайного потенциала, экранирование самосогласованным образом приводит его к критическому значению, при котором носители еще остаются свободными.

1.3. Насыщение сопротивления

Итак, выясняется, что простым наращиванием статического беспорядка не удастся создать условия для перехода Андерсона в стандартном металле и что добавление теплового беспорядка к сильному статическому может даже уменьшить сопротивление. Можно, однако, испробовать еще одну возможность — взять металл со слабым статическим беспорядком, но с большой константой электрон-фононного взаимодействия, обеспечивающей быстрый рост сопротивления с температурой, и исследовать, чего можно достичь за счет чисто теплового беспорядка. Эта попытка базируется на естественной идее об эквивалентности статического (структурного) и динамического (теплового) беспорядка с точки зрения рассеяния электронов. Скорость теплового движения ионов порядка скорости звука, т.е. порядка 10^5 см/с, в то время как фермиевская скорость электронов $v_F \approx 10^8$ см/с. Поэтому движущиеся электроны всегда видят статичную картину, включающую в себя тепловые смещения ионов.

Температурная зависимость сопротивления за счет рассеяния электронов на фононах описывается формулой Грюнрайzena

$$\begin{aligned} \rho(T) &= \rho_0 + \beta T^5, & T \ll T_D, \\ \rho(T) &= \alpha T, & T \gg T_D, \end{aligned} \quad (1.14)$$

которая при температурах выше температуры Дебая T_D выходит на линейный закон. Для достижения больших $\rho(T)$ нужны как можно большее α и высокая температура. Коэффициент α тем больше, чем сильнее электрон-фононное взаимодействие, а температура ограничивается термодинамическими процессами, например, плавлением. Этими факторами определяется выбор материалов для экспериментов.

На рис. 1.4 приведены кривые $\rho(T)$ для монокристаллов двух интерметаллидных соединений. Одно из них, Nb_3Sn , имеет температуру сверхпроводящего перехода $T_c = 18$ К и широко используется для изготовления сверхпроводящих соленоидов; второе, Nb_3Sb , тоже сверхпроводник, но у него $T_c = 0,2$ К. Сверхпроводимость здесь упоминается

не случайно. Согласно классической теории сверхпроводимости, для столь высокого T_c , как 18 К необходимо сильное электрон-фононное взаимодействие и, следовательно, можно ожидать больших значений коэффициентов β и α в формулах (1.14).

Обратимся к кривой для Nb_3Sb . Если ограничиться температурным интервалом $T < 200$ К, то мы имеем типичную кривую Грюнайзена: остаточное сопротивление при низких температурах, затем участок $T^{3,6}$ и выход на линейный рост (1.14), столь крутой, что можно было бы вполне рассчитывать на достижение критических значений при

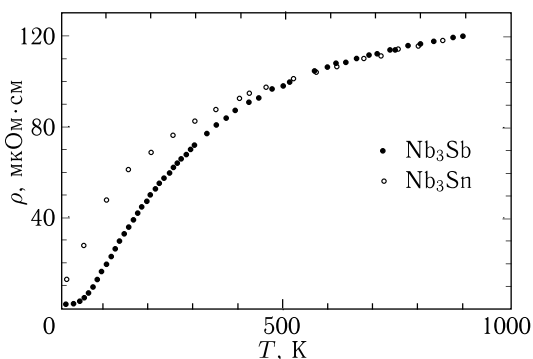


Рис. 1.4. Кривые $\rho(T)$ для монокристаллов двух интерметаллидных соединений, демонстрирующие насыщение сопротивления [4]

допустимых температурах $T \approx 900$ К. Однако вместо этого кривая $\rho(T)$ демонстрирует отчетливую тенденцию к насыщению на уровень около 150 мкОм·см. Характерно, что уровень насыщения оказался одинаковым для обоих соединений, хотя при низких температурах фононное сопротивление у Nb_3Sn гораздо выше, чем у Nb_3Sb , что коррелирует с разницей T_c .

Феноменологически экспериментальные кривые рис. 1.4 хорошо описываются моделью шунтирующего сопротивления

$$\rho^{-1} = \rho_{\text{id}}^{-1} + \rho_{\text{sh}}^{-1}; \quad (1.15)$$

когда сопротивление ρ_{id} , соответствующее формуле Грюнайзена (1.14), становится слишком большим, ток начинает течь через шунтирующее сопротивление. Эта формула успешно использовалась для описания экспериментов. Однако с ее обоснованием имеются проблемы.

Обычно формула (1.15) свидетельствует о наличии параллельных каналов проводимости, например о наличии двух независимых групп носителей с разными параметрами и законами рассеяния, находящихся в одном и том же электрическом поле. Здесь никаких оснований для существования таких групп нет. Можно придти к формуле (1.15), сохранив одну группу носителей, но введя некоторые корреляторы актов

рассеяния. Например, предположим [5], что между актами рассеяния должно пройти минимальное время τ_0 , после чего все ограничения снимаются и все происходит, как обычно. Если момент первого акта рассеяния $t = 0$, то вероятность следующего акта предполагается равной

$$p = \begin{cases} 0, & t < \tau_0, \\ 1/\tau, & t > \tau_0. \end{cases}$$

Тогда в обычную формулу для проводимости $\sigma = ne^2\tau/m$, записанную через среднее время между столкновениями τ и эффективную массу m , вместо τ войдет $\tau + \tau_0$. Отсюда получается выражение (1.15) с $\rho_{sh} = m/ne^2\tau_0$.

Применительно к рассеянию на фононах предположение о существовании времени τ_0 означает предположение о конечности времени испускания или поглощения фонона электроном, что ограничивало бы сверху частоту $1/\tau_{ph}$ величиной $1/\tau_0$. Для качественной оценки возможной величины τ_0 следует предположить, что это время, за которое электрон и фотон смещаются друг относительно друга на длину волны фонона λ . Полагая $\lambda \sim a$, получим оценку $\tau_0 \sim a/v_F \sim \hbar/\varepsilon_F$, откуда сразу следует, что

$$\rho_{sh} \sim \rho^*. \quad (1.16)$$

Хотя это как раз то, что получается в эксперименте, необходимо подчеркнуть, что предположение о конечности времени взаимодействия электрона с фононом является чисто качественным, и формулы (1.15) и (1.16) не имеют строгого теоретического обоснования.

Вместе с тем сам экспериментальный факт, что насыщение происходит при значениях сопротивления (1.16), которые типичны для высокорезистивных сплавов (см. рис. 1.3) и соответствуют предельно малой длине пробега $l \approx a$, т. е. что $\rho_{sh} \approx \rho^*$, несомненен и очень важен. Он подтверждает, что статический и динамический беспорядок воздействуют на электроны примерно одинаково. Эквивалентность двух видов беспорядка дополнительно иллюстрируется серией кривых температурной зависимости сопротивления сплава TiAl с разной концентрацией Al (рис. 1.5). Чистый титан имеет малое остаточное сопротивление и сильный температурный рост с последующей тенденцией к насыщению. Качественно кривая

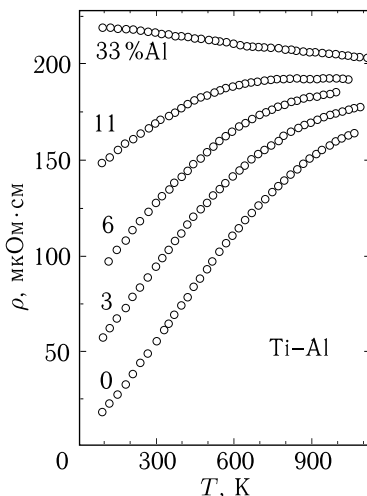


Рис. 1.5. Температурные зависимости сопротивления сплава TiAl с разной концентрацией Al [3]

$\rho(T)$ чистого титана ведет себя так же, как и на рис. 1.4. По мере увеличения концентрации Al остаточное сопротивление растет вместе со статическим беспорядком, а температурный рост уменьшается (при 33 % алюминия температурный коэффициент даже отрицательный), но предельно высокотемпературное значение сопротивления меняется сравнительно мало, оставаясь в интервале значений, типичных для высокорезистивных сплавов.

В свете существования насыщения сопротивления интересна эволюция анизотропии сопротивления в монокристаллических материалах. Ее демонстрируют кривые температурной зависимости удельного сопротивления монокристаллического иттрия на рис. 1.6. В интервале 200–300 К сопротивление и температурный коэффициент α перпенди-

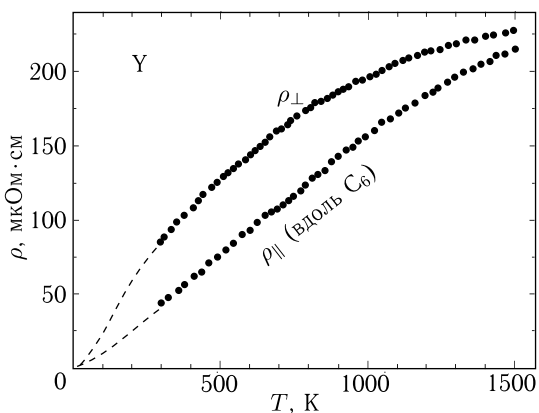


Рис. 1.6. Анизотропия температурной зависимости удельного сопротивления монокристаллического иттрия [6]

кулярно гексагональной оси примерно в 2 раза больше, чем вдоль оси. Поэтому сопротивление перпендикулярно оси приближается к критическим значениям при более низких температурах и раньше проявляет тенденцию к насыщению. В результате при 1400 К разница в сопротивлениях вдоль этих двух направлений становится уже меньше 10 %. Сопротивление насыщения ρ_{sh} в обоих направлениях, по-видимому, практически одинаковое.

Насыщение сопротивления наблюдается и в «нестандартном» металле, с несколько меньшей, чем в (1.4), концентрацией свободных зонных электронов. Обладающий металлической проводимостью окисел WO_2 имеет, согласно данным по эффекту де Гааза — ван Альфена, концентрацию носителей порядка $n \approx 2 \cdot 10^{21} \text{ см}^{-3}$. Он имеет моноклинную кристаллическую решетку и почти четырехкратную анизотропию сопротивления при комнатной температуре. На рис. 1.7 приведены кривые $\rho(T)$ для двух экстремальных направлений. В направлении, где сопротивление велико, оно демонстрирует тенденцию к насыщению

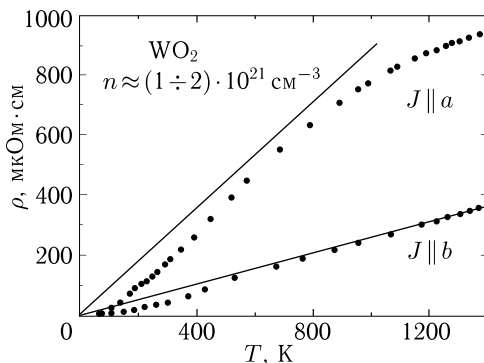


Рис. 1.7. Зависимости $\rho(T)$ для монокристалла WO_2 вдоль кристаллографических направлений с максимальным и минимальным сопротивлением [7]. Сплошные линии — прямые предельной области закона Грюнайзена (1.14). Для сопротивления при направлении тока $\mathbf{J} \parallel \mathbf{a}$ коэффициент α получен в предположении, что справедлива формула (1.15) и что $\rho_{\text{sh}} = 2000$ мкОм · см

с $\rho_{\text{sh}} \lesssim 2000$ мкОм · см. В другом направлении, где во всем температурном интервале $\rho \ll \rho_{\text{sh}}$, сопротивление следует формуле Грюнайзена без признаков насыщения.

1.4. Предел Иоффе–Регеля при большой электронной плотности

Очень поучительно свести все факты и явления, обсуждавшиеся в этой главе, на одну диаграмму. Отложим температуру T по одной оси, а частоту столкновений в энергетических единицах $\hbar/\tau_{\text{eff}}$ — по другой. В качестве характерного масштаба по обеим осям выберем фермиевскую энергию ε_{F} . Поскольку для стандартного металла $\varepsilon_{\text{F}} \sim 10000$ К, отрезок $(0 \div \varepsilon_{\text{F}})$ на температурной оси включает все интересующие нас температуры. Под частотой $1/\tau_{\text{eff}}$ подразумевается сумма частот рассеяния на статических дефектах $1/\tau$ и на фононах $1/\tau_{\text{ph}}$ (с учетом малоуглового характера фононного рассеяния при низких температурах):

$$1/\tau_{\text{eff}} = 1/\tau + 1/\tau_{\text{ph}}, \quad (1.17)$$

которая определяет статическое электросопротивление ρ . Верхний конец интервала на оси частот $\hbar/\tau_{\text{eff}} = \varepsilon_{\text{F}}$ соответствует пределу Иоффе–Регеля $l \approx k_{\text{F}}^{-1}$. Более высокие частоты столкновений попросту означали бы неправомочность модели с плоскими волнами в качестве волновых функций. Поэтому внутри квадрата $0 \leq T, \hbar/\tau_{\text{eff}} \lesssim \varepsilon_{\text{F}}$, изображенного на рис. 1.8, реализуются все транспортные явления в газе делокализованных электронов.

В этой главе обсуждался транспорт вблизи верхнего ребра этого квадрата — в области с сильным рассеянием. Сначала мы попытались

приблизиться к верхней части квадрата, двигаясь вдоль его левой стороны, при сравнительно низких температурах. Здесь на уровне $(0,1 \div 0,2)\varepsilon_F$ применима дифракционная теория жидких металлов, в существенной части принадлежащая Займану. Ее важным элементом является предположение, что волновые функции электронов — это плоские волны. Выше, вблизи левого верхнего угла квадрата, безразмерная величина $\hbar/(\varepsilon_F \tau_{\text{eff}})$ уже не является малым параметром и у теории появляются принципиальные трудности. Эксперимент устанавливает в этой области два базовых факта.

1. При сколь угодно сильном структурном беспорядке низкотемпературное сопротивление стандартного металла не превышает значения (1.6):

$$\rho \lesssim \rho^* \approx 300 \text{ мкОм} \cdot \text{см.}$$

2. Температурные зависимости сопротивления в левом верхнем углу квадрата подчиняются правилу Моойа, т. е. если сопротивление близко к значению (1.6), то рост температуры его не только не увеличивает, но даже несколько уменьшает.

Заметьте. На рис. 1.8 масштаб несколько искажен; все кривые и отмеченные пунктиром области существования растянуты вдоль оси абсцисс. Все обсуждаемые процессы происходят при

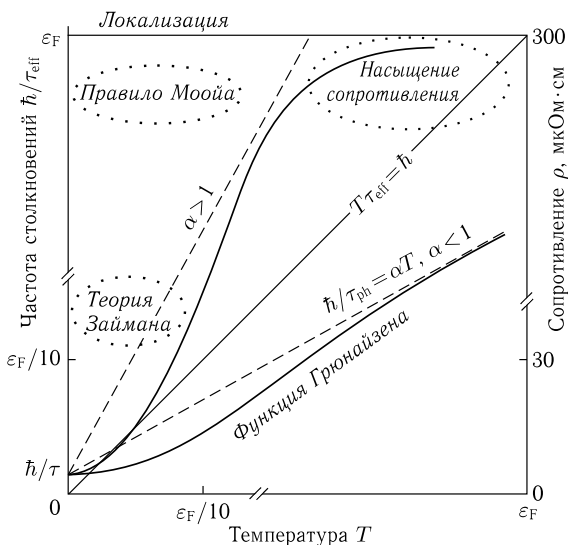


Рис. 1.8. Плоскость температура–беспорядок $(T, \hbar/\tau_{\text{eff}})$, на которой пунктирными эллипсами показаны области применимости теории Займана и правила Моойа и область, где наблюдается насыщение сопротивления. В осях $(T, \hbar/\tau_{\text{eff}})$ построены также графики сопротивления для двух случаев: (1.14) и (1.15). Штриховые прямые для обоих графиков — это высокотемпературные асимптоты функции Грюнрайзена.

температурах, не превышающих температуру плавления, т. е. при $T \lesssim \varepsilon_F/10$, в то время как частота столкновений (1.17) реально может достигать ε_F .

Переходя к попыткам пересечь верхнюю сторону квадрата за счет температурного роста сопротивления веществ с сильным электрон-фононным взаимодействием, построим диагональ квадрата

$$\hbar/\tau_{\text{eff}} = T. \quad (1.18)$$

Запомните это уравнение диагонали. Записанное в виде $\tau_{\text{eff}} = \hbar/T$, оно очень напоминает выражение для времени расфазировки двух тепловых электронов, которое определяет квантовую поправку к проводимости из-за межэлектронной интерференции. Поэтому в гл. 2 о квантовых поправках снова фигурируют и тот же квадрат, и его диагональ (см. уравнение (2.29) и рис. 2.18).

Поскольку $\rho \propto 1/\tau_{\text{eff}}$, в осях $(T, \hbar/\tau_{\text{eff}})$ можно построить функцию Грюнрайзена (1.14) (для ориентировки справа приведена также шкала в единицах удельного сопротивления). Коэффициент в высокотемпературной асимптотике функции Грюнрайзена в осях $(T, \hbar/\tau_{\text{eff}})$ становится безразмерным. Для простоты сохраним для него обозначение α . Если во всех формулах, необходимых для вычисления α , пренебрегать численными коэффициентами, то все буквенные множители сократятся и получится $\alpha = 1$. Асимптотическая прямая αT , в зависимости от численных значений параметров конкретного металла, может оказаться с любой стороны от диагонали. Если $\alpha < 1$, то асимптотическая часть кривой располагается в нижнем треугольнике $T\tau_{\text{eff}} > \hbar$. Нас больше интересовал вариант $\alpha > 1$, при котором кривая должна достигать верхней стороны квадрата. Однако, как показали эксперименты, описанные в предыдущем параграфе, в этом случае при высоких температурах появляется насыщение сопротивления, не предусмотренное формулой Грюнрайзена.

Таким образом ни увеличение статического, ни увеличение динамического беспорядка, ни их комбинация не привели к пересечению верхней стороны квадрата на рис. 1.8, за которой ожидается локализация Андерсона. Это означает, что в стандартном металле рост беспорядка сам по себе к локализации не приводит. Это экспериментальный факт, и с ним приходится считаться. Мы вернемся к обсуждению этого вопроса в гл. 5 при изложении концепции переходов металл–изолятор, а в гл. 7 покажем, как при помощи обходного маневра можно преодолеть заветную черту.

1.5. Контрольные вопросы и задачи

1. Время свободного пробега электрона при рассеянии на примесях равно τ . Сколь часто время между двумя последовательными столкновениями оказывается больше τ ; больше 2τ ; меньше $\tau/2$?

Вероятность рассеяния $dP = q dt = 1/\tau \exp(-t/\tau) dt$,

$$P(t > \tau) = 1 - \int_0^1 e^{-x} dx = 1/e = 0,37,$$

$$P(t > 2\tau) = 1 - \int_0^2 e^{-x} dx = 1/e^2 = 0,135,$$

$$P(t < \tau/2) = \int_0^{1/2} e^{-x} dx = 1 - 1/\sqrt{e} = 0,39.$$

2. Фермиевская скорость электронов $v_F = 5 \cdot 10^7$ см/с, а эффективная масса равна массе свободного электрона. Какова максимально возможная частота упругих столкновений? Что означает это ограничение и имеет ли оно отношение к соотношению неопределенности?

Из неравенства (1.2) следует $\hbar\omega_{\max} \sim mv_F^2$.

Список литературы

1. *Фабер Т.* Электронные явления переноса в жидких металлах; в *Физика металлов. 1. Электроны* / Под ред. Дж. Займана. — М.: Мир, 1972 (перевод книги *The physics of metals 1. Electrons* / Ed. J. M. Ziman. — Cambridge Univ. Press, 1969).
2. *Займан Дж.* Модели беспорядка. — М.: Мир, 1982 (перевод книги *Ziman J. M. Models of disorder*. Cambridge Univ. Press, 1979).
3. *Moos J. H.* // *Phys. Stat. Sol. (a)* **17**, 521 (1973).
4. *Fisk Z., Webb G. W.* // *Phys. Rev. Lett.* **36**, 1084 (1976).
5. *Gurvitch M.* // *Phys. Rev. B* **24**, 7404 (1981).
6. *Зиновьев В. Е., Соколов А. Л., Гельд П. В., Чуприков Г. Е., Епифанова К. И.* // *ФТТ* **17**, 3617 (1975).
7. *Гантмахер В. Ф., Кулеско Г. И., Теплинский В. М.* // *ЖЭТФ* **90**, 1421 (1986).

Глава 2

КВАНТОВЫЕ ПОПРАВКИ К ПРОВОДИМОСТИ ¹⁾

Построение теории металлов начинается с теоремы Блоха о поведении электронов в идеальной периодической решетке. Следующий этап — это учет нарушений периодичности в τ -приближении. Предполагается, что из-за нарушения периодичности электрон рассеивается, т. е. переходит из одного своего стационарного состояния в идеальной решетке в другое. Рассмотрение взаимодействия с различными типами отклонений от периодичности: примесями и другими статическими дефектами, фононами, магнонами и т. п. — наполняет параметр τ конкретным содержанием. При этом постулируется аддитивность рассеяния: вероятность $1/\tau$ перехода электрона в другое состояние есть сумма вероятностей $1/\tau_i$ на различных типах дефектов, на разных примесях, разных фононах и т. д. Предположение об аддитивности рассеяния естественно, когда акты рассеяния происходят редко, например, в чистом металле.

Сейчас мы сделаем следующий шаг и рассмотрим так называемые квантовые поправки к проводимости. Необходимым условием их появления является наличие целых серий актов рассеяния.

Заметьте. Под проводимостью мы понимаем здесь удельную проводимость с размерностью, зависящей от размерности пространства d :

$$\sigma_d [\text{Ом}^{-1} \text{см}^{2-d}], \quad d = 1, 2, 3. \quad (2.1)$$

Элементарные акты рассеяния могут быть двух типов. При одних актах энергия электрона ε_j сохраняется и, следовательно, закон изменения фазы волновой функции со временем, $\exp(i\varepsilon_j t/\hbar)$, остается прежним. Частоту таких актов рассеяния обозначим через $1/\tau$. Но есть и неупругие процессы рассеяния, например, столкновения с фононом или с другим электроном. Если в момент t_0 произошло неупругое столкновение, то из-за изменения энергии ε_j электрон «забывает» о своей фазе до столкновения, при $t < t_0$. Вероятность сбоя фазы обозначим через $1/\tau_\varphi$. Квантовые поправки, о которых пойдет речь, происходят при условии

$$\tau_\varphi \gg \tau. \quad (2.2)$$

¹⁾ Более сжатое, выделяющее лишь самое главное, обсуждение квантовых поправок имеется в книге [1]; дополнительные подробности, как экспериментальные так и теоретические, можно найти в детальных обзорах [2, 3].

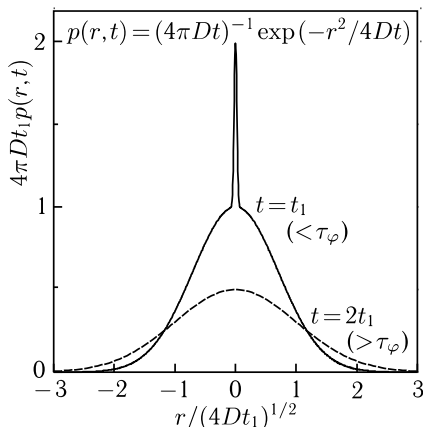
Неравенство (2.2) означает, что наиболее частыми являются акты упругого рассеяния на статическом беспорядке. На основании этого неравенства часто говорят о квантовых поправках проводимости в грязных металлах.

2.1. Слабая локализация

Слабой локализацией называется квантовая поправка к металлической проводимости, обусловленная волновыми свойствами электрона, проявляющимися на фоне диффузионного движения при большом количестве упругих рассеивателей. Электронный спектр предполагается вырожденным, $\varepsilon_F \gg T$. Электрон, находящийся в момент $t = 0$ в начале координат $r = 0$, при диффузионном движении с фермиевской скоростью v_F и средней длиной свободного пробега $l = v_F \tau$ через время $t \gg \tau$ окажется в точке $\mathbf{r}$ с вероятностью

$$p(\mathbf{r}, t) = (4\pi Dt)^{-d/2} e^{-r^2/4Dt}, \quad r^2 = \sum_i x_i^2, \quad \int p(\mathbf{r}, t) d\mathbf{r} = 1. \quad (2.3)$$

Здесь d — размерность пространства, в котором происходит диффузия, а D — коэффициент диффузии, $D = lv_F/d$. Со временем ширина распределения Δr постепенно увеличивается (рис. 2.1):



$$\Delta r \approx \sqrt{Dt} \approx l \sqrt{\frac{t}{\tau}} \approx l \sqrt{N} \quad (2.4)$$

($N = t/\tau$ — число шагов в диффузионном процессе, т. е. число упругих столкновений за время t).

Рис. 2.1. Функция $p(t)$ при диффузии на плоскости ($d=2$) в разные моменты времени t_1 и $2t_1$. Дополнительный пик на кривой $t < \tau_\varphi$ связан с квантовой интерференцией (2.5)

Формулы (2.3) и (2.4) описывают диффузию классической частицы. Если учесть волновые свойства электрона, то точка $r = 0$ окажется выделенной: в ней функция $p(r, t)$ из-за интерференции сильно изменится. Чтобы показать это, разобьем все возможные траектории, возвращающие электрон в точку $r = 0$ в момент t , на пары с одинаковым набором рассеивателей,

но с противоположными направлениями движения. Для классической частицы вероятность $p(0, t)$ есть сумма вероятностей прихода в точку $r = 0$ по разным траекториям. В квантовой механике это соответствует сложению квадратов модулей соответствующих волновых функций.

Для квантовой частицы, сохранившей к моменту $t \ll \tau_\varphi$ память об исходной фазе, складываются сами волновые функции, а вероятность $p(0, t)$ есть квадрат модуля суммы:

$$\begin{aligned} \text{без интерференции} \quad & |A_1|^2 + |A_2|^2 = 2A^2, \\ \text{с интерференцией} \quad & |A_1 + A_2|^2 = |A_1|^2 + |A_2|^2 + 2|A_1 A_2| = 4A^2. \end{aligned} \quad (2.5)$$

Поэтому вероятность найти электрон в начале координат $p(0, t)$ должна из-за интерференции удвоиться (рис. 2.1), причем ширина пика определяется соотношением неопределенности $\delta r \approx \lambda \approx 1/k_F$ (λ — де-бройлевская длина волны, k_F — фермиевский волновой век-

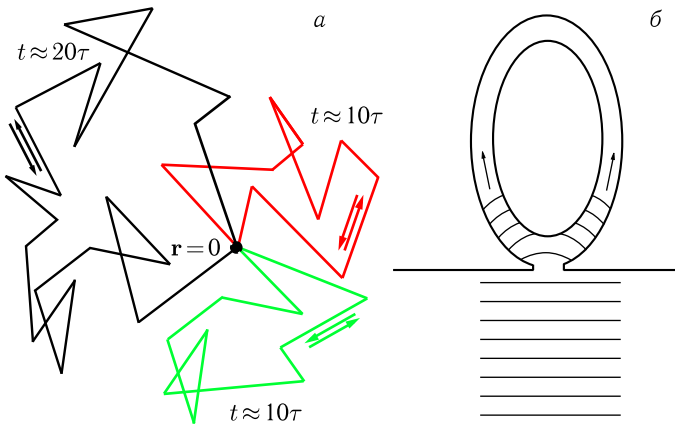


Рис. 2.2. а) Интерферирующие пары замкнутых диффузионных траекторий, вдоль которых распространяются электронные волны. б) Механическая аналогия — волны в кольцевом канале, соединенном с водоемом

тор). Разбиение на интерферирующие пары возможно только для замкнутых траекторий, т. е. тех, что заканчиваются в точке $r = 0$, в которой электрон находился в начальный момент времени (см. рис. 2.2).

Увеличение вероятности для электрона оказаться в точке $r = 0$ (т. е. по существу остаться там) и называется слабой локализацией. Оно приводит к поправкам $\delta\sigma$ в проводимости.

Пояснить сущность физических процессов, лежащих в основе слабой локализации, можно при помощи механической аналогии. Пусть кольцевой водный канал соединяется в одном месте с большим водоемом (рис. 2.2, б). Приходящая из водоема волна, разветвляясь, попадает в оба рукава канала. Если затухания волн в канале нет, то обе попавшие в рукава локальные волны обойдут весь канал и встретятся на входе.

Относительная площадь под дополнительным пиком функции $p(r, t)$ зависит от t . Поэтому для оценки относительной величины поправки к проводимости $\delta\sigma/\sigma$ нужно вычислить добавку к проводимости $d\sigma$,

возникшую за время dt из-за изменения функции $p(r, t)$, и проинтегрировать ее по времени. Рассмотрим сначала трехмерный случай $d = 3$. Объем, в любой точке которого может находиться электрон в момент времени t , порядка $(Dt)^{3/2}$. Объем, из которого он за время dt может попасть в начало координат, порядка $\lambda^2 v_F dt$. Отношение этих объемов определяет относительное количество электронов, побывавших в начале координат за время dt . Минимальное время, через которое электрон может вернуться в начало координат, это время упругого рассеяния τ . Максимальное время, через которое он, вернувшись, сможет участвовать в интерференции, это время сбоя фазы τ_φ . В результате

$$\begin{aligned} d = 3 : \quad \frac{\delta\sigma_3}{\sigma} &\approx - \int_{\tau}^{\tau_\varphi} \frac{v_F \lambda^2 dt}{(Dt)^{3/2}} \approx - \frac{v_F \lambda^2}{D^{3/2}} \left(\frac{1}{\tau^{1/2}} - \frac{1}{\tau_\varphi^{1/2}} \right) = \\ &= -(k_F l)^{-1} \left(\frac{1}{l} - \frac{1}{L_\varphi} \right). \end{aligned} \quad (2.6)$$

Введенная в (2.6) величина

$$L_\varphi \approx \sqrt{D\tau_\varphi} \approx l \sqrt{\tau_\varphi/\tau} \gg l \quad (2.7)$$

называется диффузионной длиной потери фазы. Член с L_φ в выражении (2.6) меньше члена с l , но именно через него в квантовую поправку для проводимости входит температурная зависимость, потому что τ_φ зависит от температуры и стремится к бесконечности при $T \rightarrow 0$.

Конкретная величина τ_φ и ее зависимость от температуры зависят от размерности системы и многих ее параметров, в частности, от электронной концентрации n , температуры Дебая T_D , частоты упругих столкновений $\hbar/\tau$. Мы вернемся к этому вопросу в конце этой главы.

Поскольку нормировочный множитель в $p(\mathbf{r}, t)$ в выражении (2.3) и, соответственно, знаменатель в подынтегральном выражении в (2.6) зависят от размерности d , функции $\Delta\sigma/\sigma$ для разных d оказываются совершенно различными. Характерным размером, сравнением с которым определяется размерность конкретного образца, является диффузионная длина L_φ . Пленка толщиной b и проволока диаметром b при условии

$$b \ll L_\varphi \quad (2.8)$$

являются с точки зрения диффузии объектами пониженной размерности. Вместо выражения (2.6) имеем для них

$$\begin{aligned} d = 2 : \quad \frac{\delta\sigma_2}{\sigma} &\approx - \int_{\tau}^{\tau_\varphi} \frac{v_F \lambda^2 dt}{(Dt)b} \approx \frac{v_F \lambda^2}{Db} \ln \frac{\tau}{\tau_\varphi} \approx \frac{1}{(k_F l)(k_F b)} \ln \frac{\tau}{\tau_\varphi}, \\ d = 1 : \quad \frac{\delta\sigma_1}{\sigma} &\approx - \int_{\tau}^{\tau_\varphi} \frac{v_F \lambda^2 dt}{(Dt)^{1/2} b^2} \approx \frac{2v_F \lambda^2}{Db^2} (l - L_\varphi) \approx \frac{1}{(k_F b)^2} \left(\frac{1}{l} - \frac{1}{L_\varphi} \right). \end{aligned} \quad (2.9)$$

Условие (2.8) понижения размерности очень мягкое. Оно определяет конфигурацию пространства, в котором осуществляется диффузия. При этом сам процесс диффузии может оставаться трехмерным: при $l < b$ электрон в пленке между двумя актами рассеяния может двигаться в любом направлении, в том числе и вдоль нормали к пленке; поверхность пленки в этих условиях играет роль рассеивающего центра.

Поправки $\delta\sigma_i$ в формулах (2.6) и (2.9) отрицательны. Поэтому при любой размерности проводимость σ_i , начиная с некоторой температуры, падает с ее понижением (сопротивление растет), причем эффект интерференции тем сильнее, чем ниже размерность. В трехмерном случае падение σ_3 ограничено, и σ_3 стремится к определенному пределу при $T \rightarrow 0$. При пониженной размерности поправки расходятся, так как и τ_φ , и L_φ стремятся к бесконечности при $T \rightarrow 0$. Поскольку проводимость не может быть отрицательной, должны быть ограничения применимости формул (2.6) и (2.9). Ограничением является требование относительной малости поправок:

$$\delta\sigma_i \ll \sigma \quad (i = 1, 2, 3). \quad (2.10)$$

Поучительно написать вместо относительных значений поправок к удельной проводимости (2.6) и (2.9) абсолютные значения изменений самой проводимости $\Delta\sigma_d = \delta\sigma_d b^{3-d}$, имеющих ту же размерность, что и σ_d :

$$\begin{aligned} d = 3 : \quad \Delta\sigma_3 &\approx -\text{const} + \frac{e^2}{\hbar} L_\varphi^{-1}, \\ d = 2 : \quad \Delta\sigma_2 &\approx -\frac{e^2}{\hbar} \ln \frac{\tau_\varphi}{\tau} \approx -2 \frac{e^2}{\hbar} \ln \frac{L_\varphi}{l}, \\ d = 1 : \quad \Delta\sigma_1 &\approx -\frac{e^2}{\hbar} L_\varphi. \end{aligned} \quad (2.11)$$

У них у всех одинаковый масштаб $e^2/\hbar$. Эта комбинация атомных констант, имеющая размерность обратного сопротивления, $\hbar/e^2 = 4110$ Ом, встречается во всех задачах, связанных с локализацией. Характерно, что выражения (2.11) для квантовых поправок не содержат концентрации носителей, а зависимость от τ хотя и есть, но более слабая, чем у самой проводимости. Поэтому роль интерференционных поправок возрастает по мере уменьшения собственной проводимости материала. Это еще одна причина, почему о слабой локализации говорят в связи с грязными металлами.

Заметьте. Поскольку $\tau_\varphi \rightarrow \infty$ при $T \rightarrow 0$, то при достаточно низких температурах неравенство (2.2) начнет выполняться для сколь угодно чистого металла со сколь угодно большим τ . Реально можно изготовить монокристалл чистого металла, например, индия, с длиной свободного пробега $l \approx 0,1$ см, так что параметр $k_F l$ будет иметь значение $k_F l \approx 10^6$. При $T = 0,1$ К отношение

τ_φ/τ порядка 10^3 – 10^4 , т. е. неравенство (2.2) выполняется. Однако вклад квантовой поправки в проводимость ничтожно мал:

$$\frac{\delta\sigma_3}{\sigma} \sim \frac{1}{k_F^2 L_\varphi l} \sim 10^{-14}.$$

Заметьте также. Формально даже очень толстая пленка при достаточно низкой температуре должна рассматриваться как двумерная: поскольку при $T \rightarrow 0$ длина $L_\varphi \rightarrow \infty$, ниже некоторой температуры будет выполнено неравенство (2.8). Тогда появится поправка $\Delta\sigma_2$ к классической проводимости $\sigma_3 b$.

Формулы (2.6), (2.9) и (2.11) неоднократно проверялись экспериментально, чаще всего на пленках. На рис. 2.3 сведены результаты двух разных экспериментов на пленках Си и Аи. В обоих материалах хорошо виден логарифмический рост сопротивления при температурах ниже 10 К. И хотя в одном и том же температурном интервале $1 \div 10$ К относительные величины поправок $\Delta R/R$ сильно разнятся, разница в абсолютных значениях поправок к проводимости $\Delta\sigma$ существенно меньше, как того и следовало ожидать на основании формул (2.11). Слабая зависимость величины $\Delta\sigma$ от свойств материала становится особенно наглядной, если сравнить измерения на пленках Си и Аи с измерениями на аморфных пленках In–O, у которых удельное сопротивление примерно в 1000 раз больше, а логарифмический рост начинается с 100 К. У трех пленок на рис. 2.4 значения $\Delta R/R$ составляют при десятикратном изменении температуры 15 %, 6 % и 3 %, а значения $\Delta\sigma$ лежат в области $(1 \div 2) \cdot 10^{-5} \text{ Ом}^{-1}$, будучи лишь ненамного меньше, чем на пленках Си и Аи.

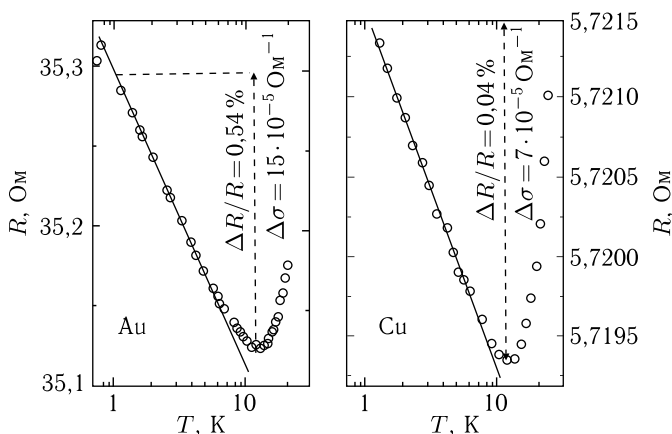


Рис. 2.3. Температурные зависимости сопротивления тонких пленок Си [4] и Аи [5]. Пунктиром показано увеличение сопротивления при десятикратном уменьшении температуры (от 10 К до 1 К).

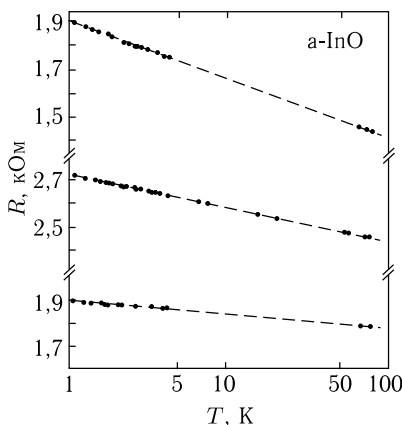


Рис. 2.4. Температурные зависимости сопротивления тонких аморфных пленок In-O [6]. Пленки различаются концентрацией кислорода, что отражается на концентрации свободных носителей

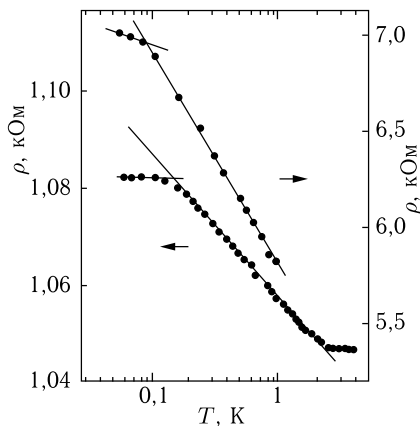


Рис. 2.5. Логарифмический температурный рост сопротивления инверсных слоев Si — данные для двух разных образцов [7]. Прямые линии служат лишь ориентирами для глаз.

Логарифмический рост сопротивления наблюдался и на истинно двумерных электронных системах, например на инверсных слоях на поверхности Si в температурном интервале 0,1–1 К (рис. 2.5). Здесь значения $\Delta R/R$ для двух образцов были 16 % и 3 %, а значения $\Delta\sigma$ порядка $2,5 \cdot 10^{-5} \text{ Ом}^{-1}$.

Заметьте. Каким бы ни был конкретный механизм неупругого рассеяния электронов, определяющий время потери фазы, величины τ_φ и L_φ являются степенными функциями температуры: $L_\varphi \propto T^s$. Поэтому на пленках квантовая поправка всегда проявляется в виде специфической зависимости $\Delta R \propto \ln T$.

Когерентное рассеяние света назад. Поскольку слабая локализация имеет волновую интерференционную природу, у нее должен быть оптический аналог. Такой аналог действительно существует: рассеяние света в мутной среде. Освещая мутную среду с одной стороны, мы можем смотреть на нее под любым углом, потому что после многократного рассеяния свет выходит с равной вероятностью во всех направлениях — феномен светлого неба. Поскольку рассеяние упруго, то модуль волнового вектора не меняется и рассеяние падающей на среду плоской световой волны $A \exp(i\mathbf{k}\mathbf{r})$ описывается расщеплением вектора $\mathbf{k}$ по поверхности сферы $|\mathbf{k}| = \text{const}$. В результате такого расщепления часть светового пучка рассеивается точно назад. Этой части соответствует на сфере точка $-\mathbf{k}$. При классическом описании эта точка ничем не

выделена, и при достаточно сильном рассеянии амплитуда рассеянной волны

$$A(\mathbf{k}) \rightarrow \text{const} \quad \text{при} \quad |\mathbf{k}| = \text{const}.$$

Для плоских волн, однако, существует интерференционный эффект, аналогичный слабой локализации. Переход из $\mathbf{k}$ в $-\mathbf{k}$ есть результат блуждания по поверхности сферы с пошаговым последовательным изменением $\mathbf{k}$ на какие-то вектора $\mathbf{q}_i$: $\mathbf{k} + \sum_i \mathbf{q}_i = -\mathbf{k}$. Существует бесчислен-

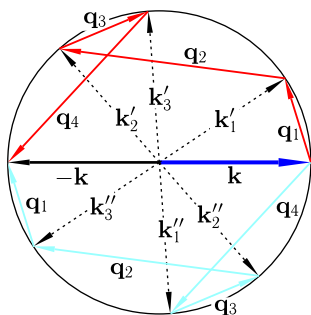


Рис. 2.6. Слабая локализация в k -представлении: для перехода $\mathbf{k} \rightarrow -\mathbf{k}$ по траектории случайного блуждания по поверхности сферы $\mathbf{k} \rightarrow \mathbf{k}'_1 \rightarrow \mathbf{k}'_2 \rightarrow \mathbf{k}'_3 \rightarrow -\mathbf{k}$ есть парная траектория $\mathbf{k} \rightarrow \mathbf{k}''_1 \rightarrow \mathbf{k}''_2 \rightarrow \mathbf{k}''_3 \rightarrow -\mathbf{k}$, интерференция с которой приводит к усилению рассеяния назад

ное количество последовательностей $\{\mathbf{q}_i\}$, и их все тоже можно разбить на интерферирующие пары, состоящие из одного и того же набора векторов, но включающихся во взаимно обратном порядке — см. рис. 2.6 (частичная перестановка, например, двух случайных векторов в сумме по i не годится, потому что промежуточные значения $\mathbf{k}_i$ окажутся не на поверхности сферы). Для каждой интерферирующей пары траекторий в k -пространстве можно повторить рассуждения (2.5), из которых следует относительное увеличение интенсивности рассеяния назад.

Экспериментальная установка для реализации такой «слабой локализации» схематически изображена на рис. 2.7. Регистрируется угловая зависимость интенсивности света, рассе-

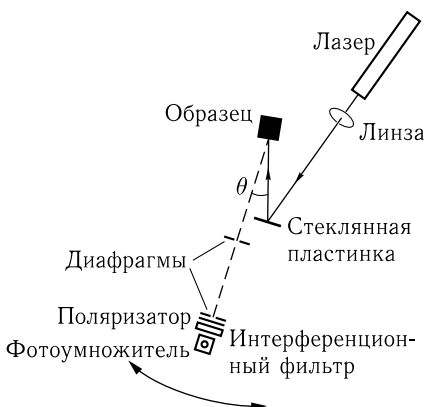


Рис. 2.7. Когерентное рассеяние света назад — схема эксперимента [8]

янного взвесью в воде полистироловых шариков диаметром 0,46 мкм, занимающих в общей сложности 10 % объема. Результаты представлены на рис. 2.8 (кривая *a*) вместе с аналогичным сигналом от кюветы с чистой водой (кривая *б*) и от пустой кюветы (кривая *в*).

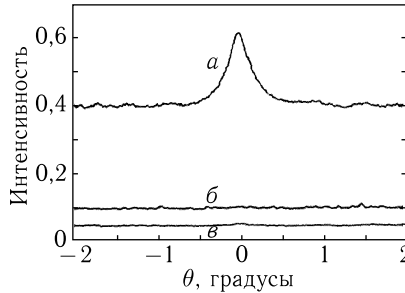


Рис. 2.8. Экспериментальные кривые, демонстрирующие существование когерентного рассеяния света назад [8]. *a* — сигнал, отраженный от кюветы с рассеивающими микрочастицами, взвешенными в воде; *б* — от кюветы с чистой водой; *в* — от пустой кюветы

2.2. Влияние магнитного поля на слабую локализацию

Магнитное поле B изгибает электронную траекторию между двумя последовательно рассеивающими электрон центрами, причем изгибает в разные стороны при движении в противоположных направлениях. Поэтому углы рассеяния на каждой примеси будут различаться на величину порядка $l/R = \Omega\tau$ (R — циклотронный радиус, а $\Omega = eB/mc$ — циклотронная частота). Но поскольку обычно речь идет о грязном металле с не очень большим τ , то в не очень большом поле $\Omega\tau \ll 1$ разницей углов можно пренебречь по сравнению с другим, гораздо более сильным эффектом.

Когда электрон проходит по замкнутому контуру в магнитном поле, то у его волновой функции Ψ появляется дополнительный фазовый множитель:

$$\Psi \rightarrow \Psi \exp\left(i \frac{e}{\hbar c} \int \mathbf{A} d\mathbf{l}\right) = \Psi \exp\left(\pm \frac{i\pi BS}{\Phi_0}\right). \quad (2.12)$$

Здесь $\mathbf{A}$ — вектор-потенциал магнитного поля, $\Phi_0 = \pi\hbar c/e$ — квант магнитного потока, а $BS = \Phi$ — магнитный поток через замкнутый контур электронной траектории. Знак в показателе экспоненты зависит от того, проходит электрон этот контур по или против часовой стрелки. Поскольку электрон движется по парам интерферирующих траекторий в противоположных направлениях, при возвращении его в начало координат на двух траекториях появится разность фаз $\varphi = 2\pi(\Phi/\Phi_0)$.

Площадь S проекции замкнутого контура на плоскость, перпендикулярную полю, порядка квадрата среднего расстояния $r(t)$, на которое электрон уйдет от начала координат за время t . Поэтому

$$\varphi = 2\pi \frac{BS}{\Phi_0} \approx \frac{Br^2}{\Phi_0} \approx \frac{BDt}{\Phi_0}. \quad (2.13)$$

Наличие разности фаз означает, что вместо второго соотношения (2.5) имеет место

$$|A_1 + A_2|^2 = |A_1|^2 + |A_2|^2 + 2|A_1||A_2|\cos\varphi = 2A^2(1 + \cos\varphi). \quad (2.14)$$

Пока φ мало, это уточнение несущественно. Но при $\varphi > 1$ становится важным разброс площадей S пар интерферирующих траекторий по сравнению с оценкой (2.13). На рис. 2.2 это разница площадей контуров, нарисованных штриховой и пунктирной линиями. Среднее значение $\overline{\cos\varphi} = 0$, так что магнитное поле разрушает интерференционную добавку. Чем больше время t и средняя площадь $\bar{S} \approx Dt$ сформировавшихся за это время замкнутых траекторий, тем меньшее поле B разрушает интерференционный вклад. Времена $t > \tau_\varphi$ нас не интересуют, потому что интерференция на этих временах все равно разрушена неупругими столкновениями. А окрестность верхнего предела интегрирования τ_φ в интегралах (2.6) и (2.9) существенна: именно она входит через L_φ в формулы (2.11). Поэтому, положив в (2.13) слева $\varphi = 1$, а справа $t = \tau_\varphi$, получим значение поля, которое уже начинает разрушать слабую локализацию:

$$B_\varphi = \frac{c\hbar}{e} (D\tau_\varphi)^{-1}. \quad (2.15)$$

Записав коэффициент диффузии в виде $D \approx \varepsilon_F \tau / m$ и введя циклотронную частоту $\Omega_\varphi = eB_\varphi / mc$ вместо поля B_φ , получим из соотношения (2.15)

$$\Omega_\varphi \tau \frac{\varepsilon_F \tau_\varphi}{\hbar} \approx 1, \quad \text{откуда} \quad \Omega_\varphi \tau \ll 1. \quad (2.16)$$

Неравенство (2.16) подтверждает, что в поле B_φ можно пренебречь искривлением электронных траекторий.

Введем магнитное время

$$\tau_B = \frac{\Phi_0}{BD} \approx \frac{r_B^2}{D} \approx \Omega^{-1} (k_F l)^{-1}, \quad (2.17)$$

где $r_B = (\hbar c / 2eB)^{1/2}$ — магнитная длина (часто употребляется также «магнитный радиус»). Когда $\tau_B \ll \tau_\varphi$, т. е. когда $B \gg B_\varphi$, в интегралах

(2.6) и (2.9) нужно заменить верхний предел: $\int_\tau^{\tau_\varphi} \rightarrow \int_\tau^{\tau_B}$.

Тогда

$$0 < \Delta\sigma(B) - \Delta\sigma(0) \approx \begin{cases} 2 \frac{e^2}{\hbar} \ln \frac{L_\varphi}{r_B}, & d = 2, \\ \frac{e^2}{\hbar} \left(\frac{1}{r_B} - \frac{1}{L_\varphi} \right), & d = 3 \end{cases} \quad (l \ll r_B \leq L_\varphi). \quad (2.18)$$

В уравнении (2.18) добавка к проводимости $\Delta\sigma(B)$ сравнивается с добавкой в нулевом поле $\Delta\sigma(0)$. Иногда при анализе кривых удобнее в качестве базы выбрать не $\sigma(0)$, а проводимость в сильном поле.

Рис. 2.9 иллюстрирует разрушение магнитным полем слабой локализации в пленках Mg. Интегральное падение сопротивления тем больше, чем ниже температура T . Это следует и из формулы (2.9),

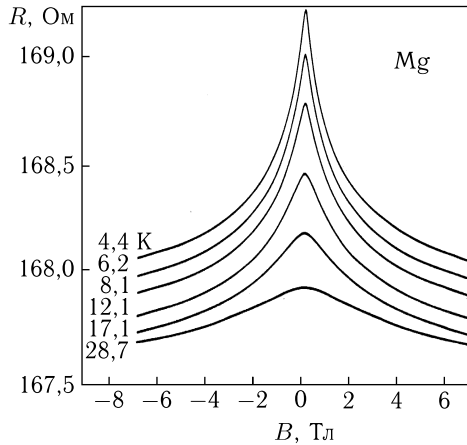


Рис. 2.9. Магнетосопротивление тонкой холодноосажденной пленки Mg при разных температурах [2]

где при уменьшении T растет верхний предел интегрирования. При этом область основного падения сопротивления смещается по мере уменьшения T в меньшие поля. Интерференция практически полностью разрушена в поле B_{el} ,

$$B_{el} = \frac{\hbar c}{e l^2}, \quad \text{когда} \quad \tau_B = \tau, \quad (2.19)$$

т.е. когда магнитное время равно упругому. Если постепенно уменьшать сильное поле, то сначала включается интерференция на траекториях с самой маленькой площадью S проекции на плоскость, перпендикулярную магнитному полю, $S \sim l^2$. Это произойдет в поле B_{el} . По мере дальнейшего уменьшения поля включается интерференция на траекториях со все большими значениями S . Это сопровождается

ростом сопротивления. Рост прекращается в поле B_φ ,

$$B_\varphi = \frac{\hbar c}{eL_\varphi^2}, \quad \text{когда} \quad \tau_B = \tau_\varphi. \quad (2.20)$$

Из этих рассуждений следует, что температура должна влиять на кривые $\Delta\sigma(B)$ лишь в малых полях $B \gtrsim B_\varphi$, а в больших полях $B \approx B_{el}$ квантовые поправки от температуры зависеть не должны. Наблюдаемый на рис. 2.9 вертикальный сдвиг кривых $R(B)$ друг относительно друга в сильных полях объясняется температурной зависимостью классической части проводимости в магнитном поле.

Таким образом магнитное поле можно использовать как инструмент для обнаружения слабой локализации: свидетельством ее существования является отрицательное магнетосопротивление в слабом поле. Это экспериментальное доказательство существования слабой локализации как бы методом от противного, потому что мы видим процесс ее разрушения. Но как демонстрация слабой локализации зависимости от поля на рис. 2.9 ничуть не менее убедительны, чем температурные зависимости на рис. 2.3–2.5.

Из рассуждений, сопровождающих формулу (2.14), следует, что интерференционная добавка разрушается полем порядка B_φ из-за разброса площадей замкнутых интерферирующих траекторий. Если сделать так, чтобы все замкнутые электронные траектории имели одну и ту же площадь S проекции на плоскость, перпендикулярную магнитному полю, то интерференционный вклад не разрушится, а начнет осциллировать с полем по мере роста φ с периодом

$$\Delta B = \frac{\Phi_0}{S} = \frac{\pi \hbar c}{eS}. \quad (2.21)$$

Такая конфигурация была реализована экспериментально. Если на кварцевую нить диаметром $2r \approx 1 \div 2$ мкм напылить тонкий слой $\delta r \ll r$ металла, то получится цилиндрическая пленка. Площадь S проекции замкнутой диффузионной траектории на плоскость, перпендикулярную оси цилиндра, равна либо нулю, либо πr^2 . Магнитное поле, направленное вдоль оси, на интерференцию траекторий с $S = 0$ влияния не оказывает. В противоположность этому вклад от траекторий с $S = \pi r^2$ в проводимость вдоль оси цилиндра должен осциллировать с периодом (2.21). Для того чтобы этот вклад был сравним по величине с обычной квантовой поправкой, требуется выполнение условия $S < D\tau_\varphi \approx L_\varphi^2$: тонкий цилиндр и низкая температура. Однако осциллирующую добавку удастся, по-видимому, наблюдать и в условиях $S \gtrsim L_\varphi^2$.

На рис. 2.10 собраны экспериментальные кривые, полученные двумя экспериментальными группами на Mg, Li и Al. Как видно из

графиков, с теоретическими расчетами удается согласовать не только периоды осцилляций, но и зависимость их амплитуды от поля и монотонный ход.

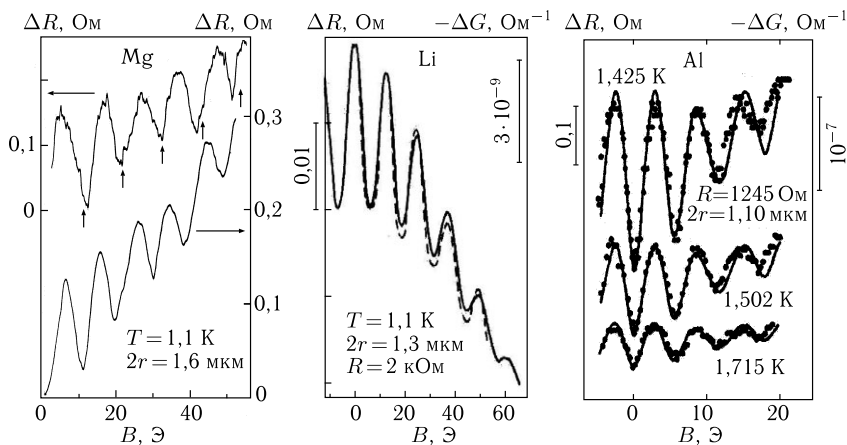


Рис. 2.10. Обусловленные слабой локализацией осцилляции в продольном магнитном поле сопротивления цилиндрических пленок Mg (из работы [9], здесь сплошные линии — это экспериментальные кривые для двух разных пленок, а стрелками показаны рассчитанные значения поля, при которых должны наблюдаться экстремумы сопротивления), Li (из работы [10], эксперимент показан сплошной линией, а расчет — штриховой) и Al (из работы [11] — экспериментальные точки и расчетные кривые). На графиках для Li и Al указан масштаб изменений не только сопротивления ΔR , но и проводимости $\Delta G = \Delta R/R^2$.

Из наших предыдущих рассуждений следует, что в поле $B = 0$ сопротивление должно иметь максимум. Однако, так происходит только на Li, а на Mg и Al осцилляции имеют противоположный знак. Мы вернемся к этому факту в следующем параграфе после обсуждения влияния на интерференцию спин-орбитального взаимодействия.

Цилиндрическая форма пленки выделяет определенную группу диффузионных траекторий и сводит усреднение по всем возможным площадям проекций S к сумме вкладов от траекторий с площадями $S = 0$ и $S = \pi r^2$. Неполноценное усреднение возможно не только за счет особой формы образца; оно может возникнуть также и в образце произвольной формы, но достаточно малого размера $L \lesssim L_\varphi$. При изменении магнитного поля вклад в проводимость от каждой замкнутой диффузионной траектории осциллирует со своим периодом. Количество замкнутых траекторий в малом образце ограничено, поэтому полного усреднения их осциллирующих вкладов с периодами (2.21) не происходит, и при изменении поля возникают так называемые *универсальные флуктуации кондактанса*.

Обычно универсальные флуктуации кондактанса наблюдают при измерениях транспортных свойств тонких проволок, в которых поперечные размеры много меньше диффузионной длины L_φ . На рис. 2.11 показано расположение нескольких контактов на проволоке из Sb сечением $0,12 \times 0,08$ мкм². В принципе, для измерения сопротивления любой из них можно использовать и как токовый, и как потенциальный. Обозначение $R_{mn;j \rightarrow k}$ означает, что измеряется разность потенциалов $V_m - V_n$ между контактами m и n при токе $J_{j \rightarrow k}$, пропускаемом через контакты j и k , т. е. $R_{mn;j \rightarrow k} = (V_m - V_n)/J_{j \rightarrow k}$.

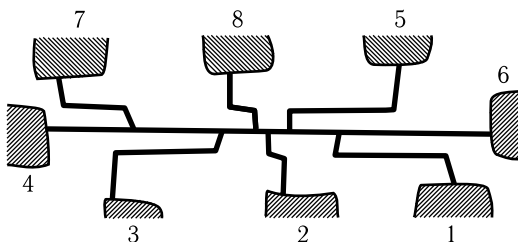


Рис. 2.11. Расположение контактов на проволоке Sb сечением $0,12 \times 0,08$ мкм², изготовленной для изучения универсальных флуктуаций кондактанса [12]

На рис. 2.12, а приведены результаты измерения сопротивления $R_{83;6 \rightarrow 4}$ по обычной четырехконтактной схеме на участке цепи длиной $0,66$ мкм между контактами 8 и 3 [12]. Амплитуда осцилляций кондактанса $\Delta\sigma$ должна иметь тот же масштаб $e^2/\hbar$, что и сам эффект слабой локализации, определяемый уравнениями (2.11). Поэтому естественным масштабом для осцилляций сопротивления является величина

$$\Delta R_\varphi = R_\varphi^2 \Delta Y_\varphi = R_\varphi^2 (e^2/\hbar),$$

где R_φ и Y_φ соответственно — сопротивление и кондактанс проволоки длиной L_φ , так что $Y_\varphi = 1/R_\varphi$. В эксперименте на рис. 2.12, а длина $L_\varphi \approx 1$ мкм, а средняя амплитуда осцилляций $\Delta R/\Delta R_\varphi \approx 0,4 \div 0,5$.

Универсальные флуктуации кондактанса обладают специфическим свойством нелокальности: они приводят к осцилляциям напряжения не только на активном участке цепи, через который идет ток, но и на окрестных пассивных участках цепи, находящихся от активного участка не слишком далеко, на расстояниях $L < L_\varphi$. Это эффектно демонстрирует измерение сопротивления $R_{86;4 \rightarrow 3}$ (рис. 2.12, б). С классической точки зрения разность потенциалов на контактах 8 и 6 должна тождественно равняться нулю. Однако, к нулю близко лишь среднее значение напряжения $\bar{V}_8(B) - \bar{V}_6(B)$, а амплитуда флуктуаций того же порядка, что и на рис. 2.12, а: $\Delta R/\Delta R_\varphi \approx 0,2 \div 0,3$. Нелокальные флуктуации напряжения $V_8 - V_6$ возникают из-за того, что на активном участке проволоки между контактными площадками 4 и 3 имеется неравновесная добавка к функции распределения. Входящие в нее

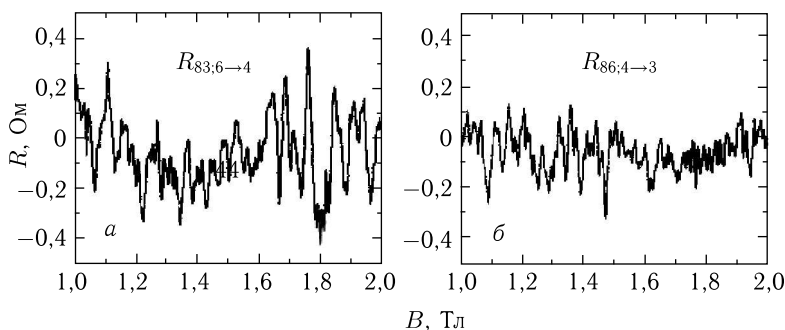


Рис. 2.12. Локальные (а) и нелокальные (б) флуктуации сопротивления на проволоке Sb, изображенной на рис. 2.11, при изменении магнитного поля. Температура $T = 0,048$ К, ток $J_{j \rightarrow k} = 20$ нА [12]

электроны, диффундируя вдоль проволоки, проходят, в частности, замкнутые диффузионные траектории и приобретают в результате слабой локализации добавку к волновой функции. Она имеет характерный размер L_φ , т. е. захватывает также и участок между контактами 8 и 6. Электронная плотность на этом участке отклоняется от равновесной, что вызывает появление электрического поля и разности потенциалов.

К вопросу о терминологии и классификации. Обсуждая универсальные флуктуации кондактанса, мы столкнулись с типичной мезоскопической задачей. Мезоскопика это раздел физики твердого тела, занимающийся системами, в которых число частиц недостаточно для использованием стандартных методов статистики, и флуктуации характеризующих систему величин порядка их средних значений [13]. В мезоскопике принципиальным является сравнение размера системы L с длиной сбоя фазы L_φ : в системах с $L < L_\varphi$ рассмотрение должно быть принципиально квантовомеханическим, так как необходимо учитывать интерференцию электронных волн. Поэтому все аспекты слабой локализации очень важны для мезоскопики. По существу теория квантовых поправок к проводимости является мостиком между мезоскопикой и физической макросистем.

2.3. Антилокализация

Две электронные волны, движущиеся в противоположных направлениях по замкнутой диффузионной траектории и интерферирующие на выходе, имеют фиксированную проекцию спина $+1/2$ или $-1/2$. Фактически предполагается, что по каждой замкнутой траектории движутся два электрона с разными спинами, каждый в виде пары

электронных волн, распространяющихся в противоположных направлениях. Каждый электрон интерферирует сам с собой, когда две волны встречаются в исходной точке. Если переворотов спина нет, то эти два электрона совершенно независимы и вносят аддитивный вклад в квантовую поправку к проводимости. Однако все меняется при наличии спин-орбитального взаимодействия, когда спин электрона может переворачиваться при упругом рассеянии. Тогда движения по двум траекториям в одном направлении перемешиваются.

Спин-орбитальное взаимодействие возникает благодаря тому, что магнитный момент μ , движущийся со скоростью $\mathbf{v}$, создает в неподвижной системе координат электрическое поле $\mathbf{e} \propto \mu \times \mathbf{v}$. Это поле взаимодействует с полем $\mathbf{E}$ зарядов, входящих в ионы, что и приводит к перевороту спина. Величина поля E , а вместе с ней и интенсивность спин-орбитального рассеяния, которая пропорциональна

$$\mathbf{Ee} \propto \mathbf{E} [\mu \times \mathbf{v}], \quad (2.22)$$

сильно зависят от суммарного количества зарядов в ионах, т.е. от атомного номера Z атомов: $\tau_{so}^{-1} \propto (Z\alpha)^4$, где $\alpha = 1/137$ – постоянная тонкой структуры. Поэтому влияние спин-орбитального взаимодействия сильнее в материалах с тяжелыми элементами.

Поскольку далеко не каждый акт упругого рассеяния сопровождается переворотом спина, время переворота спина много больше времени упругого рассеяния: $\tau_{so} \gg \tau$. Отношение τ_{so}/τ тем больше, чем слабее спин-орбитальное взаимодействие. При этом τ_{so} не зависит от температуры. Поскольку $\tau_{\varphi} \rightarrow \infty$ при $T \rightarrow 0$, то при достаточно низких температурах

$$\tau \ll \tau_{so} \ll \tau_{\varphi}. \quad (2.23)$$

Чтобы учесть перемешивание двух одинаковых замкнутых диффузионных траекторий с разными спинами, удобно считать, что в каждую сторону по траектории движется пара электронов. Спин такой пары на входе и на выходе может быть в одном из четырех состояний: полный спин либо 0 (синглет), либо 1, причем во втором случае возможны три разные проекции (триплет). Волновая функция пары Ψ имеет соответственно четыре компоненты:

$$\Psi = \begin{pmatrix} \Psi_0 \\ \Psi_{1,-1} \\ \Psi_{1,0} \\ \Psi_{1,1} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} (\varphi_+^{(1)} \varphi_-^{(2)} - \varphi_-^{(1)} \varphi_+^{(2)}) \\ \varphi_-^{(1)} \varphi_-^{(2)} \\ \frac{1}{\sqrt{2}} (\varphi_+^{(1)} \varphi_-^{(2)} + \varphi_-^{(1)} \varphi_+^{(2)}) \\ \varphi_+^{(1)} \varphi_+^{(2)} \end{pmatrix}, \quad (2.24)$$

где $\varphi^{(1)}$ и $\varphi^{(2)}$ – волновые функции отдельно первого и второго электронов, а нижние индексы + или – относятся к спиновой части

волновой функции и означают разные проекции спинов. Такая пара интерферирует сама с собой за счет обхода замкнутой траектории в противоположных направлениях. Полный интерференционный член возникает как сумма четырех слагаемых — от трех волновых функций Ψ_{1m} с полным спином 1 (с проекциями $m = 0, \pm 1$) и функции Ψ_0 со спином 0. Опуская выкладки, выпишем и прокомментируем окончательный результат

$$\frac{\delta\sigma_d}{\sigma} \approx - \int_{\tau}^{\tau_{\varphi}} \frac{v\lambda^2 dt}{b^{3-d}(Dt)^{d/2}} \left(\frac{3}{2} e^{-t/\tau_{so}} - \frac{1}{2} \right), \quad d = 1, 2, 3. \quad (2.25)$$

По сравнению с формулами (2.6) и (2.9) в подынтегральном выражении в (2.25) появился новый множитель в виде скобки, состоящей из двух слагаемых. Первое слагаемое возникло от трех триплетных состояний Ψ_{1m} со спином 1. Экспоненциальное затухание в нем означает, что интерференция этих функций существенна только до тех пор, пока $t < \tau_{so}$, т.е. пока электроны помнят об изначальном спине. В этом интервале можно считать, что $\exp(-t/\tau_{so}) \approx 1$, так что выражение в скобках порядка единицы, и подынтегральное выражение совпадает с тем, которое было без учета спина. Это естественно: ведь переворота спина еще не было.

Самое интересное происходит на временах

$$\tau_{so} < t < \tau_{\varphi}, \quad (2.26)$$

когда первым слагаемым в скобках можно пренебречь и остается только второе отрицательное слагаемое, происходящее от интерференции волновой функции Ψ_0 . Знак минус означает, что оставшаяся после переворотов спина интерференционная добавка не уменьшает, а увеличивает проводимость. Окончательный знак поправки зависит от того, какой интервал, $t < \tau_{so}$ или $t > \tau_{so}$, дает больший вклад в интеграл (2.25). При $\tau_{so} \ll \tau_{\varphi}$ вклад интерференции в проводимость положителен. Иногда это называют антилокализацией.

Не следует думать, что *фазы* волновой функции синглетной пары после обхода замкнутого контура в противоположных направлениях отличаются на π . Интерференция фиксируется на временах $t > \tau_{so}$, когда уже произошли перевороты практически всех спинов. Каждый переворот спина превращает триплетную пару в синглетную или наоборот, синглетную пару в триплетную. Поэтому при $t > \tau_{so}$ синглетная часть волновой функции уже на три четверти состоит из «бывших» триплетов. А синглетные и триплетные волновые функции устроены по-разному: как видно из формулы (2.24), триплетные компоненты симметричны относительно перестановок частиц, а синглетная — антисимметрична. Именно это приводит к появлению минуса при вычислении *амплитуды* интерференционного вклада от синглетной части волновой функции.

Обратите внимание. Все сказанное справедливо только для спин-орбитального взаимодействия и не относится к рассеянию с переворотом спина на парамагнитной примеси. Перевернув спин примесного центра, электрон со спином μ оказывается в состоянии со спином $-\mu$, которое ранее было пустым. Такой процесс аналогичен процессу обычного рассеяния без переворота спина и приводит не к антилокализации, а к слабой локализации.

Конечно, все эти рассуждения не заменяют прямых вычислений и потому могут вызывать чувство неудовлетворения. Посмотрим, однако, какой вопрос в проблеме антилокализации является чисто техническим, а какой — принципиальным. Хорошо известно, что интерференция может приводить не только к увеличению, но и к уменьшению амплитуды волны, причем конкретный знак изменений часто не очевиден. Достаточно вспомнить дискуссию в Парижской Академии Наук в связи с формулами Френеля: тот факт, что из них следует наличие светлого пятна на экране на оси пучка, диффрагирующего на круглом непрозрачном диске, сочли поначалу свидетельством несостоятельности теории волновой природы света, а после экспериментального обнаружения такого пятна — наоборот, убедительным доказательством ее правильности. Слабая антилокализация является столь же логически возможным результатом интерференции электронных волн, как и слабая локализация. Принципиальным является другой факт: электронная волна продолжает принимать участие в интерференции после того, как спин-орбитальное взаимодействие перевернуло спин электрона.

Электроны находятся в контакте с окружающей их внешней квантовой системой, которую можно назвать термостатом. Судить о том, как скажется акт рассеяния на когерентности электронной волны, т. е. на ее способности к интерференции, можно по состоянию термостата. Упругое рассеяние бесспинового электрона не оставляет следов в термостате. Поэтому когерентность сохраняется, хотя электрон при рассеянии может оказаться в любом состоянии на сфере $|\mathbf{k}| = \text{const}$. При неупругом рассеянии электрона на решетке состояние термостата меняется: в нем появляется или исчезает фоновый фонон. Соответственно электрон теряет когерентность.

Переворот спина тоже может произойти двумя способами. Когда электрон рассеивается с переворотом спина на магнитной примеси, то из-за сохранения полного спина на примеси остается перевернутый спин. Это означает изменение состояния термостата, хотя вообще говоря, не меняются ни энергия электрона, ни энергия термостата. И действительно, рассеяние на магнитных примесях вносит наряду с фоновым рассеянием вклад в $1/\tau_\varphi$ и *не приводит* к антилокализации. Переворот спина при рассеянии на немагнитных примесях в результате спин-орбитального взаимодействия не оставляет следов в термостате — и сохраняет когерентность электрона. Результат интерференции тем не менее меняется, потому что изменились интерференционные условия.

Множество когерентных состояний заполняет теперь поверхности двух сфер: $|\mathbf{k}_-| = \text{const}$ и $|\mathbf{k}_+| = \text{const}$, где индексы $+$ или $-$ означают направление проекции спина.

Антилокализацию удобнее всего наблюдать по зависимостям от разрушающей интерференцию магнитного поля. На рис. 2.13 приведены зависимости $R(B)$ на пленках Cu. Атомы Cu более тяжелые, чем

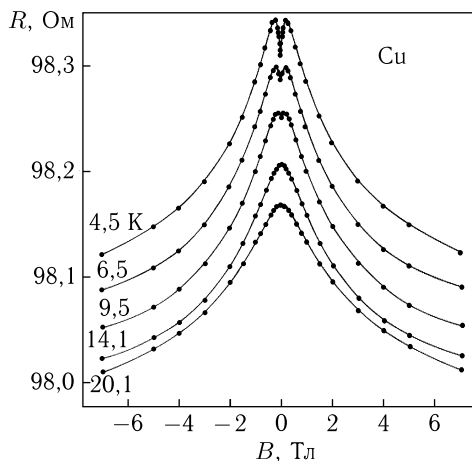


Рис. 2.13. Магнетосопротивление тонкой холодноосажденной пленки Cu при разных температурах [2]. Подбором параметров удается идеально провести теоретические кривые через экспериментальные точки

атомы Mg, и спин-орбитальное взаимодействие при рассеянии должно проявляться сильнее. И действительно, сравнивая с соответствующими кривыми для Mg на рис. 2.9, видим при самых низких температурах появление дополнительного минимума в слабых полях. Согласно (2.17), магнитное время $\tau_B \propto B^{-1}$. Поэтому при увеличении поля от нуля τ_B сравнивается сначала с τ_φ и съедает ту часть интеграла (2.25), которая ответственна за антилокализацию. При этом проводимость должна падать, а сопротивление расти. Затем τ_B сравнивается с τ_{s0} , область антилокализации в подынтегральной функции исчезает, и дальнейшее увеличение поля разрушает обычную слабую локализацию. Поскольку τ_{s0} не зависит от температуры, а τ_φ с ростом температуры падает, при более высоких T изначально нет области (2.26) и, соответственно, нет минимума на кривых $R(B)$.

Интересно проследить за эволюцией кривых $R(B)$ по мере уменьшения времени τ_{s0} . Такую возможность дает эксперимент на пленках Mg с напыленным сверху небольшим количеством атомов Au, играющих роль центров, рассеивающих с переворотом спина. Около кривых на рис. 2.14 справа указано количество Au в процентах заполнения

моноатомного слоя, а слева — отношение времен τ_φ/τ_{so} , полученное как подгоночный параметр при сравнении экспериментальных кривых с теоретическими.

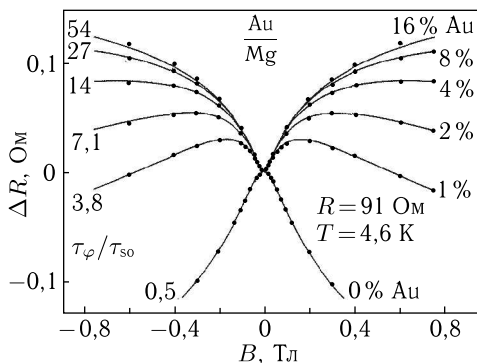


Рис. 2.14. Магнетосопротивление тонких холодноосажденных пленок Mg, покрытых золотом. Толщина покрытия в процентах от одного атомного слоя указана справа от кривых. Слева указано отношение параметров τ_φ/τ_{so} , использованное для построения теоретических кривых, проходящих через экспериментальные точки; это отношение меняется практически пропорционально толщине покрытия золотом [2]

Вернемся к осцилляциям квантовой поправки к проводимости цилиндрических пленок (рис. 2.10). Обратную фазу осцилляций на Mg и Al по сравнению с Li естественно объяснить влиянием спин-орбитального взаимодействия. Хотя все три металла относятся к легким элементам, все же Li имеет существенно меньший атомный номер: $Z_{Li} = 3$, $Z_{Mg} = 12$, $Z_{Al} = 13$, так что различие в τ_{so} при прочих равных условиях превышает два порядка.

В формулу (2.22), качественно описывающую спин-орбитальное взаимодействие, кристаллический потенциал U входит дважды: через электрическое поле, $\mathbf{E} = \nabla U$, и через скорость, которая есть производная от закона дисперсии, $\mathbf{v} = \hbar^{-1} \partial \varepsilon / \partial \mathbf{k}$, и которая тоже определяется кристаллическим полем. Поэтому кристаллическая структура материала и, в частности, его симметрия являются вторым, помимо атомного номера Z , существенным фактором, определяющим спин-орбитальное взаимодействие и, соответственно, рассеяние. Особенно важным является центр инверсии кристаллического поля. При отсутствии центра инверсии выражение (2.22) формирует второй вклад в спин-орбитальное взаимодействие, который в ряде случаев может оказаться основным.

Большинство приведенных выше классических примеров слабой локализации и антилокализации реализованы на металлических пленках. Сейчас исследования двумерного электронного газа в основном осуществляются на гетероструктурах, в которых электроны находятся

в узкой квантовой яме на границе двух кристаллических сред. При этом обычно используются вещества, кристаллические структуры которых не имеют центра инверсии. Однако спин-орбитальное взаимодействие в таком двумерном газе зависит также от формы и от асимметрии самой ямы (эффект Рашбы–Бычкова):

$$\nabla U = (\nabla U)_{\text{cryst}} + (\nabla U)_{\text{heter}}. \quad (2.27)$$

А параметры ямы обычно можно менять, подавая электрическое напряжение на расположенный параллельно яме затвор. Поэтому появляется возможность переходить от локализации к антилокализации и обратно в двумерном электронном газе, меняя напряжение V_g на затворе. На рис. 2.15 и 2.16 приведены результаты двух подобных экспериментов.

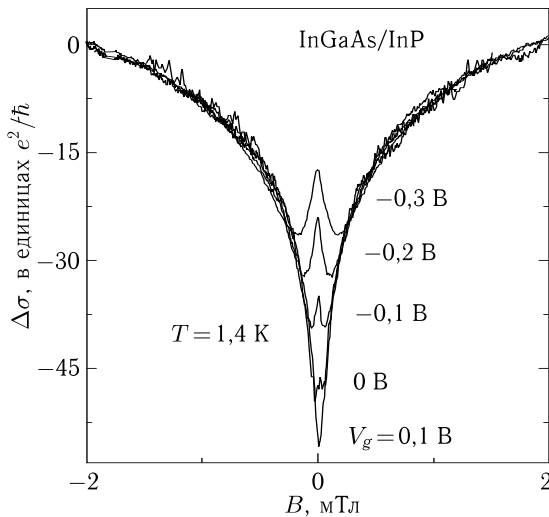


Рис. 2.15. Антилокализация, регулируемая напряжением V_g на затворе, в квантовой яме InGaAs/InP с высокой подвижностью электронов [14]. Температура 1,4 К. Кривые совмещены при значении поля 2 мТл. Предполагается, что это поле полностью разрушило интерференционную поправку ($\Delta\sigma = 0$)

Не случайно, что на этих рисунках вдоль оси ординат отложена именно проводимость, которая определяется формулами (2.11) и (2.25), а не сопротивление. Иначе было бы трудно сравнивать между собой кривые при разных напряжениях V_g на затворе, поскольку V_g сильно меняет концентрацию носителей и сопротивление двумерного электронного газа, а $\Delta\rho = \Delta\sigma/R^2$. Такого сильного различия сопротивлений не было в экспериментах, представленных на рис. 2.9, 2.13 и 2.14. Поэтому там приведены графики сопротивления, которое непосредственно измеряется в экспериментах.

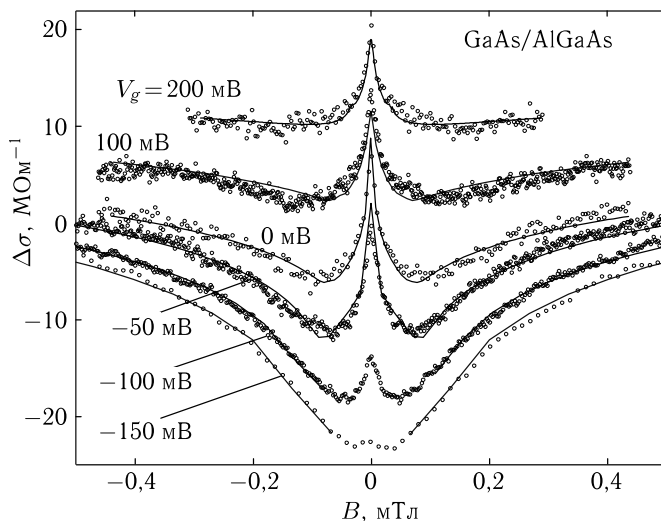


Рис. 2.16. Антилокализация, регулируемая напряжением V_g на затворе, в квантовой яме GaAs/AlGaAs с высокой подвижностью [15]. Точками показаны полученные экспериментальные данные. Температура 300 мК. Сплошные кривые это результат подгонки теоретических формул с использованием трех параметров. Кривые смещены для ясности вдоль вертикальной оси

2.4. Межэлектронная интерференция ¹⁾

К вопросу о терминологии.

Название этого параграфа означает, что в нем пойдет речь о явлениях, вытекающих в конечном счете из волновых свойств электронов. По то же самое можно сказать и о слабой локализации. Поэтому иногда, говоря об электронной интерференции, имеют в виду именно ее. А вместо межэлектронной интерференции говорят об электрон-электронном взаимодействии или пользуются названием эффект Аронова–Альтшулера. Однако, если придерживаться оптической терминологии, то слабую локализацию следует считать результатом электронной дифракции, а эффект Аронова–Альтшулера — результатом интерференции.

Слабая локализация происходит в результате интерференции электрона с самим собой благодаря возможности его движения по разным траекториям. Но интерференция возможна также и при взаимодействии разных электронов. Электроны, оказавшиеся вблизи друг друга в момент $t = 0$, при отсутствии рассеивателей разлетаются по баллистическим траекториям с фермиевской скоростью v_F , и расстоя-

¹⁾ Последовательное изложение теории см. в [3].

ние r между ними линейно растет со временем, $r \propto v_F t$. В условиях диффузии среднее расстояние между ними растет гораздо медленнее: $r \propto l\sqrt{t/\tau} \propto v_F\sqrt{t\tau}$. Диффузия удерживает электроны вблизи друг друга и тем самым создает условия для интерференции.

При слабой локализации время сбоя фазы τ_φ определяется неупругими столкновениями. При межэлектронном взаимодействии механизм сбоя фазы иной: он связан с тем, что электроны изначально имеют несколько разные энергии ε_i . Изменение фазы каждого электрона со временем определяется соотношением $\exp(i\varphi) = \exp[i(\varepsilon_i/\hbar)t]$. Поэтому одинаковые в момент $t = 0$ фазы двух электронов с разностью энергий $\Delta\varepsilon$ через время $\hbar/\Delta\varepsilon$ будут уже различаться на величину порядка единицы.

Упругое рассеяние существенно только для электронов из интервала энергий

$$\varepsilon_F - T \lesssim \varepsilon_i \lesssim \varepsilon_F + T. \quad (2.28)$$

Поэтому среднее значение $\Delta\varepsilon \sim T$, а характерное время расфазировки

$$\tau_{ee} \approx \hbar/T. \quad (2.29)$$

Соответствующий размер области интерференции равен

$$L_{ee} \approx l\sqrt{\frac{\tau_{ee}}{\tau}} \approx v_F\sqrt{\frac{\hbar\tau}{T}} \approx \sqrt{\frac{\hbar D}{T}}. \quad (2.30)$$

Величины τ_{ee} и L_{ee} играют в межэлектронной интерференции ту же роль, что величины τ_φ и L_φ в слабой локализации, хотя механизм влияния на проводимость этих двух эффектов принципиально различный.

Диффузионный характер движения в первую очередь меняет частоту межэлектронных соударений. Когда две заряженные частицы пролетают друг мимо друга, энергия и импульс у каждой из них после разлета оказываются иными, нежели до встречи. В баллистическом пределе средний передаваемый при столкновении импульс — порядка k_F , эффективный размер области взаимодействия — $1/k_F$, а эффективное время взаимодействия — $\hbar/\varepsilon_F$. Поскольку каждый акт рассеяния должен удовлетворять принципу Паули и закону сохранения энергии, начальное и конечное состояния рассеивателя могут лежать только в интервале энергий (2.28). Температура ограничивает число возможных столкновений: число электронов, с которыми может столкнуться данный электрон, пропорционально T , и число конечных состояний, в которых может оказаться этот электрон в результате каждого такого столкновения, тоже пропорционально T . Это вносит в вероятность рассеяния $\hbar/\tau_e$ множитель T^2 , так что частота столкновений в баллистическом режиме ($T\tau/\hbar \gg 1$)

$$\hbar/\tau_e^{\text{ball}} \sim T^2/\varepsilon_F. \quad (2.31)$$

Здесь множитель $1/\varepsilon_F$ добавлен из соображений размерности, поскольку ε_F — это единственный независимый параметр размерности энергии.

Выражение для частоты столкновений (2.31) принципиально меняется, если электроны не пролетают друг мимо друга по баллистическим траекториям, а диффундируют вблизи друг друга, рассеиваясь на примесях. Эффективный размер области взаимодействия теперь $L_{ee} \gg 1/k_F$ велик, а переданный в результате этого взаимодействия импульс $q \approx 1/L_{ee}$ мал. Именно это ограничение на величину передаваемого импульса, а не закон сохранения энергии, определяет частоту столкновений τ_e в этой области. Вместо T^2 в τ_e входит q^d . Чтобы получить размерность энергии для переменной $\hbar/\tau_e$, величину $q^d \propto 1/L_{ee}^d$ следует умножить на плотность состояний на ферми-уровне, тоже зависящую от размерности пространства d : $g_d = \partial n / \partial \varepsilon \sim \varepsilon_F^{d/2-1} m^{d/2}$ (где $m = \hbar k_F / v_F$ это масса электрона). Вместо выражения (2.31) получится

$$\hbar/\tau_e^{\text{diff}} \sim \frac{1}{g_d L_{ee}^d}. \quad (2.32)$$

Отсюда для частоты столкновений в диффузионном режиме ($T\tau/\hbar \ll 1$)

$$\begin{aligned} \hbar/\tau_e^{\text{diff}} &\sim T^{d/2} \varepsilon_F^{1-d} \tau^{-d/2} = \\ &= \begin{cases} T^{3/2} \varepsilon_F^{-2} (\tau/\hbar)^{-3/2} = \left(\frac{T^2}{\varepsilon_F} \right) (T\tau/\hbar)^{-1/2} (\varepsilon_F \tau/\hbar)^{-1}, & d = 3, \\ T \varepsilon_F^{-1} (\tau/\hbar)^{-1} = \left(\frac{T^2}{\varepsilon_F} \right) (T\tau/\hbar)^{-1}, & d = 2, \\ T^{1/2} (\tau/\hbar)^{-1/2} = \left(\frac{T^2}{\varepsilon_F} \right) (T\tau/\hbar)^{-3/2} (\varepsilon_F \tau/\hbar), & d = 1. \end{cases} \quad (2.33) \end{aligned}$$

Заметьте. Диффузионный характер движения электронов на временах $\hbar/T$ не отменяет рассеяние с большой передачей импульса (2.31), а добавляет к нему рассеяние (2.33) в диффузионном канале.

Столкновения (2.33) вносят непосредственный вклад в скорость релаксации энергии и в формирование времени сбоя фазы τ_φ . Но обмен импульсом между электронами при таких столкновениях несущественен, и сами по себе они не вносят заметного вклада в сопротивление, поскольку по смыслу самой постановки задачи $1/\tau_e \ll 1/\tau$. Но параллельно с включением диффузионного канала рассеяния взаимодействие электронов в процессе диффузии приводит и к изменению электронного спектра вблизи ферми-уровня.

Рассмотрим взаимодействие двух электронов с энергиями, отсчитанными от ферми-уровня, ε и $-\varepsilon$. Время их расфазировки расходится при $\varepsilon \rightarrow 0$: $\tau_{ee} = \hbar/\varepsilon$. В гамильтониане их взаимодействия есть обменный член. Он появляется благодаря тому, что ввиду тождественности электронов, их можно менять местами. Этот член не влияет на рассеяние. Он лишь вносит небольшие добавки в энергию ε и плотность

состояний $g(\varepsilon) = dn/d\varepsilon$ электронов. Обычно это не имеет значения, поскольку эти добавки не зависят от энергии. Но поправка в обменный член от взаимодействия при диффузионном движении не теряется на общем фоне, потому что соответствующий вклад в плотность состояний зависит от энергии и максимален на ферми-уровне: эффективное время взаимодействия электронов с энергиями ε и $-\varepsilon$ тем больше, чем меньше $|\varepsilon|$. При $T = 0$ плотность состояний вблизи ферми-уровня в зависимости от энергии ε имеет вид

$$\Delta g(T = 0, \varepsilon) \sim (\hbar D)^{-d/2} \cdot \begin{cases} \sqrt{|\varepsilon|}, & d = 3, \\ \ln(\varepsilon\tau/\hbar), & d = 2, \\ 1/\sqrt{|\varepsilon|}, & d = 1. \end{cases} \quad (2.34)$$

Это главный результат межэлектронного взаимодействия в диффузионном канале. Удерживая электроны друг около друга в $\mathbf{r}$ -пространстве, диффузия тем самым увеличивает эффективное время их взаимодействия, что приводит к расталкиванию уровней.

Вообще говоря, взаимодействие влияет на само понятие плотности состояний, потому что при наличии взаимодействия распределение электронов по энергии зависит от их количества. Из-за этого плотность состояний, входящая, например, в формулы для теплоемкости, отличается от той, которая определяет транспортные свойства и вероятность туннелирования и о которой здесь идет речь. Этот вопрос более подробно обсуждается в Приложении Б. В частности, там отмечается, что кулоновское взаимодействие приводит к уменьшению эффективной туннельной плотности состояний из-за конечности времени рассасывания пространственной неоднородности зарядов, которая возникает при туннелировании. Из сказанного здесь видно, как диффузия влияет на это время.

В результате взаимодействия у функции $g(T = 0, \varepsilon)$ при всех d появляется особенность на ферми-уровне, $\varepsilon = \varepsilon_F$. Конечная температура, увеличивая среднюю разность энергий взаимодействующих электронов, обрезает особенность в плотности состояний:

$$g(T, \varepsilon) \approx \begin{cases} g_F, & |\varepsilon_F - \varepsilon| > \hbar/\tau, \\ g(0, \varepsilon), & T < |\varepsilon_F - \varepsilon| < \hbar/\tau, \\ g(0, \varepsilon = T), & |\varepsilon_F - \varepsilon| < T. \end{cases} \quad (2.35)$$

Это схематически изображено на рис. 2.17, где особенность у функции $g(0, \varepsilon)$ показана пунктиром. Именно то, что в непосредственной окрестности ферми-уровня плотность состояний зависит от температуры, приводит к температурной зависимости соответствующей поправки к проводимости. Оценить величину этой поправки можно так же, как это было сделано для слабой локализации.

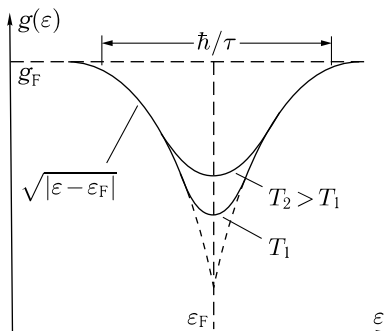


Рис. 2.17. Минимум плотности состояний на ферми-поверхности, появляющийся вследствие межэлектронного взаимодействия, и его зависимость от температуры (трехмерный случай)

Для того чтобы произошла межэлектронная интерференция, электроны, бывшие вблизи друг друга в момент $t = 0$, должны повстречаться вторично в течение времени τ_{ee} . Вероятность такой встречи η описывается теми же самыми интегралами, которые фигурировали в теории слабой локализации:

$$\eta(\Delta\varepsilon) \approx \int_{\tau}^{\hbar/\Delta\varepsilon} \frac{v\lambda^2 dt}{b^{3-d}(Dt)^{d/2}}, \quad d = 1, 2, 3, \quad (2.36)$$

где $\Delta\varepsilon$ — разность энергий двух интерферирующих электронов. Это предопределяет функциональное подобие формул квантовых поправок к проводимости от слабой локализации и от межэлектронного взаимодействия:

$$\begin{aligned} d = 3 : \quad \Delta_{ee}\sigma_3 &\approx -\text{const} + \frac{e^2}{\hbar} L_{ee}^{-1}, \\ d = 2 : \quad \Delta_{ee}\sigma_2 &\approx -\frac{e^2}{\hbar} \ln \frac{\tau_{ee}}{\tau} \approx -2 \frac{e^2}{\hbar} \ln \frac{L_{ee}}{l}, \\ d = 1 : \quad \Delta_{ee}\sigma_1 &\approx -\frac{e^2}{\hbar} L_{ee}. \end{aligned} \quad (2.37)$$

Эти поправки и поправки (2.11) от слабой локализации имеют одинаковую зависимость от размерности d , те же критерии размерности (2.8) и т. д. Однако вместо длины L_φ , определенной соотношениями (2.7), в них входит длина расфазировки (2.30).

Естественным способом экспериментального наблюдения межэлектронной интерференции должно было бы быть изучение транспортных свойств грязных металлов. Однако проблема в том, чтобы различить межэлектронную интерференцию и слабую локализацию. Главная зацепка — это отсутствие у межэлектронной интерференции сильной зависимости от магнитного поля. Однако использовать эту зацепку быва-

ет не очень просто. Поэтому основные прямые демонстрации этого эффекта получены путем изучения плотности состояний. Минимум плотности состояний можно увидеть непосредственно в специальных туннельных экспериментах, описанных в Приложении Б.

2.5. Сравнительный анализ квантовых поправок

Формулы для квантовых поправок: (2.11), описывающие слабую локализацию (электронную дифракцию), и (2.37), описывающие эффект Аронова–Альтшулера (межэлектронную интерференцию), — очень похожи. Разница между ними заключена в функциях $\tau_\varphi(T)$ и $\tau_{ee}(T)$, которые следует проанализировать.

Построим все относящиеся к делу характерные обратные времена $\hbar/\tau_i$ и частоты рассеяния на одной плоскости с осями $x \equiv T/\varepsilon_F$ и $y \equiv \gamma_i = \hbar/\tau_i \varepsilon_F$ нормированными на фермиевскую энергию ε_F (см. рис. 2.18). Частота упругих столкновений $\hbar/\tau$ представляет собой

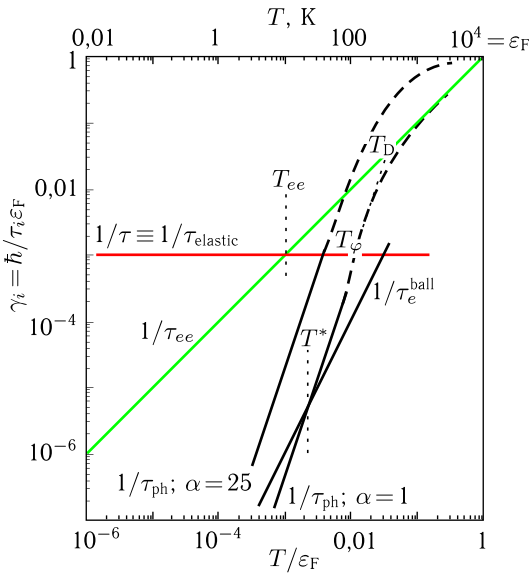


Рис. 2.18. Температурные зависимости обратных времен и частот рассеяния, определяющих квантовые поправки к проводимости в трехмерных системах. Около прямых проставлены значения индекса i при τ , определяющие характер соответствующего конкретного процесса

на этом графике горизонтальную прямую, а обратное время расфазировки (2.29) — прямую, выходящую из начала координат под углом 45° . Межэлектронная интерференция реализуется слева от точки

пересечения этих прямых при

$$T < T_{ee} \equiv \hbar/\tau. \quad (2.38)$$

Обратите внимание. Нарисованная на рис. 2.18 область изменений T и τ представляет собой часть квадрата, который использовался по другому поводу в гл. 1 — см. там рис. 1.8. Отличие лишь в том, что на рис. 2.18 используются логарифмические шкалы, так что степенные функции имеют вид прямых линий.

Чтобы сделать все соотношения на рис. 2.18 конкретными, положим фермиевскую энергию $\varepsilon_F = 10^4$ К (это позволяет построить наверху вторую шкалу непосредственно в градусах), температуру Дебая $T_D = 400$ К, а частоту упругих столкновений $\hbar/\tau\varepsilon_F = (k_F l)^{-1} = 10^{-3}$. Тогда получается $T_{ee} = 10$ К.

Построим также на плоскости $[T, 1/\tau_i]$ частоту неупругих столкновений с фононами

$$\hbar/\tau_{ph} \sim \begin{cases} \alpha' T^3/T_D^2, & T \ll T_D, \\ \alpha T, & T \gg T_D. \end{cases} \quad (2.39)$$

Это — полная частота электрон-фононных столкновений, которая при $T \ll T_D$ отличается от транспортной частоты (1.14) из-за того, что при низких температурах рассеяние электронов на фононах становится малоугловым. Коэффициент обычно порядка единицы. Низкотемпературная зависимость (2.39) отличается от зависимости при межэлектронном рассеянии (2.31), так что при достаточно низких температурах $T < T^* \equiv T_D^2/\varepsilon_F$ межэлектронное рассеяние становится основным, и именно оно определяет τ_φ и L_φ :

$$1/\tau_\varphi = \max \{ 1/\tau_e^{\text{ball}}, 1/\tau_{ph} \}. \quad (2.40)$$

В нашем конкретном примере $T^* = 16$ К, а граница области слабой локализации $T_\varphi \approx 120$ К $> T_{ee}$ находится правее, чем граница области межэлектронной интерференции T_{ee} . Такая ситуация допустима: когерентность одноэлектронной конфигурации может сохраняться там, где многоэлектронная когерентность уже нарушена.

Формула (2.39) содержит, помимо букв, численный коэффициент α . Как мы знаем из гл. 1 (параграф 1.3), существуют материалы с большой константой электрон-фононного взаимодействия, т.е. с большим численным значением α . Это означает параллельный сдвиг кривой τ_{ph} вверх. На рис. 2.18, помимо кривой $1/\tau_{ph}$, соответствующей $\alpha = 1$, приведена кривая $\alpha = 25$. В результате роста α в низких температурах расширяется область, где время сохранения фазовой когерентности лимитируется рассеянием на фононах, T_φ смещается в более низкие температуры, а при высоких температурах появляется насыщение сопротивления.

Увеличение частоты неупругого рассеяния $\hbar/\tau$ (рост беспорядка) приводит к росту T_{ee} , т.е. к увеличению температурного интервала,

в котором реализуется межэлектронная интерференция. При этом, как видно из рис. 2.18, этот интервал может захватить даже комнатные температуры. Тогда в электросопротивлении появляются два зависящих от температуры и сравнимых по величине вклада разных знаков: растущий с температурой вклад из-за рассеяния на фононах и падающая с температурой квантовая поправка. Добавив к классическому выражению для проводимости трехмерной среды выражение для квантовой поправки $\delta\sigma = (e^2/\hbar)(1/L_\varphi)$, получим:

$$\sigma = \sigma_{cl} + \delta\sigma = \frac{e^2}{\hbar} k_F \left[\frac{1}{\gamma + \gamma_{ph}} + (\gamma\gamma_{ph})^{1/2} \right]. \quad (2.41)$$

В этом выражении зависящая от температуры функция γ_{ph} стоит в первом слагаемом в знаменателе, а во втором в числителе, причем относительный вес первого (классического) слагаемого уменьшается по мере роста частоты статического рассеяния γ . На этом пути можно попытаться объяснить правило Мооя (параграф 1.2 в гл. 1).

Следует, однако, иметь в виду, что с точки зрения теории частота γ должна быть существенно меньше единицы, так как вся теория построена в первом приближении теории возмущений.

Все сказанное выше о τ_φ , строго говоря, относится к трехмерным системам. В системах с меньшей размерностью основной причиной сбоя электронной фазы в области температур $T < T_{ee}$ оказывается межэлектронное рассеяние (2.33) в диффузионном канале. Расчет τ_φ при этом осложняется тем, что рассеяние (2.33) сопровождается малой передачей энергии $\delta\varepsilon \ll T$, из-за чего $\tau_\varphi \neq \tau_e^{\text{diff}}$. Результат многочисленных актов межэлектронного рассеяния в диффузионном режиме можно представить как диффузию в энергетическом пространстве (ср. с формулой (2.4))

$$\Delta\varepsilon(t) = \delta\varepsilon(t/\tau_e^{\text{diff}})^{1/2}, \quad (2.42)$$

Определив τ_φ как время, через которое дополнительный набег фазы $\Delta\varphi$ станет порядка единицы, получим

$$\Delta\varphi(t = \tau_\varphi) \sim \tau_\varphi \frac{\Delta\varepsilon(\tau_\varphi)}{\hbar} = \tau_\varphi \frac{\delta\varepsilon}{\hbar} (\tau_\varphi/\tau_e^{\text{diff}})^{1/2} \sim 1, \quad \tau_\varphi \sim \left(\frac{\tau_e^{\text{diff}} \hbar^2}{(\delta\varepsilon)^2} \right)^{1/3}. \quad (2.43)$$

Как бы то ни было, в системах пониженной размерности ($d = 2, 1$) τ_φ и, соответственно, область существования квантовых поправок формируется процессами межэлектронного рассеяния.

2.6. Контрольные вопросы и задачи

1. Длины L_φ (формула (2.7)) и L_{ee} (формула (2.30)) связаны с определением размера квантовой когерентности. Имеет ли одна из них приоритет по отношению к другой в том смысле, что всегда ли в формулы (2.11) нужно подставлять L_φ , а в формулы (2.37) — L_{ee} ?

Если по формулам (2.7), (2.30) окажется, что $L_\varphi < L_{ee}$, то в (2.37) следует использовать L_φ .

2. Может ли пленка толщиной b в некотором интервале температур считаться двумерной с точки зрения эффекта слабой локализации, но трехмерной с точки зрения эффекта Аронова–Альтшулера или наоборот?

Возможно, если $L_{ee} < b < L_\varphi$.

3. Сравнить добавку от слабой локализации в зависимости проводимости от поля B , если

- а) поле нормально к пленке;
- б) поле лежит в плоскости пленки;
- в) пленка свернута в трубку диаметром q с осью вдоль поля.

4. Сколь бы малыми ни были остаточное сопротивление ρ_0 и $1/\tau$, т. е. сколь бы чистым ни был металл, при достаточно низкой температуре $T < \hbar/\tau$ он достигает грязного, диффузионного, предела, так что в его проводимости формально должна быть учтена квантовая поправка. Какова ее относительная величина в системах разной размерности?

$$d = 3, \quad \Delta\sigma/\sigma \sim (k_F l)^{-2} l/L_\varphi;$$

$$d = 2, \quad \Delta\sigma/\sigma \sim \nu^{-1} (k_F l)^{-1} \ln l/L_\varphi;$$

$$d = 1, \quad \Delta\sigma/\sigma \sim \nu^{-2} L_\varphi/l.$$

Число каналов $\nu = k_F b$ от чистоты материала не зависит.

5. Почему в экспериментах Бергмана на Mg (рис. 2.9) наблюдается слабая локализация (в нулевом поле производная $dR/dB < 0$), а в экспериментах Шарвина на Mg (рис. 2.10) — антилокализация ($dR/dB > 0$)?

При использовании формулы (2.25) для описания эксперимента Шарвина в качестве нижнего предела в интеграле следует взять время диффузии вдоль периметра цилиндра, а верхнего предела — бесконечность.

6. При увеличении с 1,5 мкм до 1,7 мкм диаметра кварцевой нити, на которую напыляли цилиндрическую пленку, амплитуда осцилляций от электронной дифракции уменьшилась в 2 раза. Оценить, исходя из этих данных, длину сбоя фазы.

См. предыдущую задачу.

7. При температуре T , много меньше дебаевской, диффузионная длина, определяемая рассеянием на акустических фонах, равна L_φ . Во сколько раз увеличится L_φ , когда температура уменьшится в 10 раз? Тот же вопрос для длины L_{ee} .

L_φ увеличится в $T^{3/2}$ раз, а L_{ee} — в $T^{1/2}$ раз.

8. Покажите, что при $d = 3$ вклад диффузионного канала (2.33) в общую частоту e – e -рассеяния всегда мал по сравнению со вкладом баллистического канала (2.31), хотя последний убывает с температурой как T^2 , а первый — лишь как $T^{3/2}$.

Существует граница по температуре, ниже которой диффузионный канал не работает.

Список литературы

1. *Абрикосов А. А.* Основы теории металлов. — М.: Физматлит, 2006.
2. *Bergmann G.* Weak localization in thin films // *Phys. Rep.* **107**, 1 (1984).
3. *Altshuler B. I., Aronov A. G.* Electron-electron interaction in disordered conductors, in *Electron-electron interactions in disordered systems*. (Eds.: A. L. Efros, M. Pollak). — North-Holland, 1985.
4. *Van der Dries L., Van Haesendonck C., Bruynseraede Y., Deutscher G.* // *Phys. Rev. Lett.* **46**, 565 (1981).
5. *Дорожкин С. И., Долгополов В. Т.* // *Письма в ЖЭТФ* **36**, 15 (1982).
6. *Ovadyahu Z., Imry Y.* // *Phys. Rev. B* **24**, 7440 (1981).
7. *Bishop D. J., Tsui D. C., Dynes R. C.* *Phys. Rev. Lett.* **44**, 1153 (1980).
8. *Wolf P., Maret G.* // *Phys. Rev. Lett.* **55**, 2696 (1985).
9. *Шарвин Д. Ю., Шарвин Ю. В.* // *Письма в ЖЭТФ* **34**, 285 (1981).
10. *Альтшулер Б. Л., Аронов А. Г., Спивак Б. З., Шарвин Д. Ю., Шарвин Ю. В.* // *Письма в ЖЭТФ* **35**, 476 (1982).
11. *Gijs M., Van Haesendonck C., Bruynseraede Y.* // *Phys. Rev. B* **30**, 2964 (1984).
12. *Haucke H., Washburn S., Benoit A. D., Umbach C. P., Webb R. A.* // *Phys. Rev. B* **41**, 12454 (1990).
13. *Имри Й.* Введение в мезоскопическую физику. — М.: Физматлит, 2002 (перевод книги *Imry Y.* *Introduction to Mesoscopic Physics*. — Oxford University Press, 1997).
14. *Studenikin S. A., Coleridge P. T., Poole P., Sachrajda A.* // *Письма в ЖЭТФ* **77**, 362 (2003).
15. *Miller J. B., Zumbuhl D. M., Marcus C. M., Lyanda-Geller Y. B., Goldhaber-Gordon D., Campman K., Gossard A. C.* // *Phys. Rev. Lett.* **90**, 076807 (2003).

Глава 3

ВЛИЯНИЕ МЕЖЭЛЕКТРОННОГО ВЗАИМОДЕЙСТВИЯ НА ЭЛЕКТРОННЫЙ ЭНЕРГЕТИЧЕСКИЙ СПЕКТР ¹⁾

3.1. Переход Пайерлса ²⁾

Исходный потенциал, в котором находятся все электроны в твердом теле, создается ионными остовами. Однако неравномерное распределение электронов в пространстве приводит к появлению дополнительных электрических полей и к перенормировке исходного потенциала. Поэтому каждый отдельный электрон чувствует потенциал, созданный не только ионами, но и остальными электронами. В частности, такой перенормировкой является экранирование. В металлах, где концентрация делокализованных электронов очень велика, а экранировка очень сильна, от исходного потенциала остается только малая часть, называемая псевдопотенциалом. Псевдопотенциал воздействует на одно-электронный спектр только благодаря сохранившейся трансляционной симметрии исходного потенциала.

Однако часто воздействие электронов на собственный энергетический спектр не ограничивается простым экранированием. Электроны могут существенно изменить затравочный потенциал, даже понизить его симметрию. Самый яркий пример — это неустойчивость Пайерлса. Рассмотрим регулярную одномерную цепочку атомов с периодом a . Зоной Бриллюэна для нее является отрезок на оси k от $-\pi/a$ до π/a (см. рис.3.1). На этом отрезке имеется $1/a$ состояний в расчете на единицу длины цепочки, в каждом из которых может находиться два электрона, всего $2/a$ мест для электронов. Пусть каждый атом отдает в общее пользование один электрон, так что их концентрация $n = 1/a$. Предположим для простоты, что температура $T = 0$. Тогда n электронов займут $n/2$ состояний с наименьшей энергией в центре зоны Бриллюэна, расположенные на отрезке от $-\pi/2a$ до $\pi/2a$. Предположим, что период атомов цепочки каким-то образом удвоился. Тогда период в обратном пространстве вдвое уменьшился и как раз на месте раздела между свободными и занятыми состояниями появилась дополнительная граница зоны Бриллюэна. На границе зоны в спектре $\varepsilon(k)$

¹⁾ Большая часть материала этой главы излагается также в книге [1].

²⁾ Более подробное изложение можно найти в книгах [2] и [3].

имеется щель, а производная $d\varepsilon/dk$ равна нулю с обеих сторон. Поэтому, как видно из рис. 3.1, энергия свободных состояний $|k| > \pi/2a$ при появлении дополнительной границы повысилась, а энергия занятых

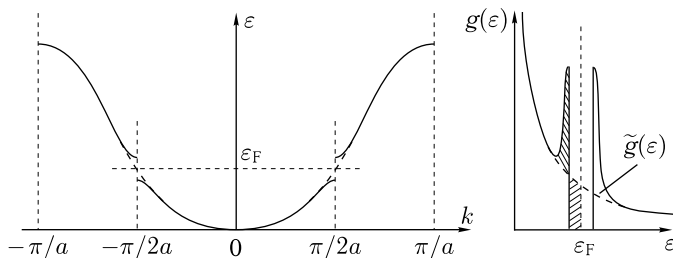


Рис. 3.1. Появление щели вокруг ферми-уровня в одномерной зоне Бриллюэна для цепочки металлических атомов. Справа плотность состояний $g(\varepsilon)$; две заштрихованные площади равны между собой. Пунктиром показана плотность состояний $\tilde{g}(\varepsilon)$ невозмущенного одномерного электронного газа

состояний $|k| < \pi/2a$ понизилась, а вместе с этим понизилась и энергия всей электронной системы. Положение ферми-уровня ε_F относительно дна зоны в точке $k = 0$ при этом не изменилось.

Таким образом, для электронной системы удвоение периода цепочки выгодно, потому что приводит к уменьшению ее энергии. Вопрос в том, как этого удвоения периода добиться. Простейший способ — это малое смещение $(\Delta a)_2$ каждого второго атома; хотя при этом возникает проигрыш в упругой энергии, которая определяла период a , малое смещение $(\Delta a)_2$ обычно оказывается все же выгодным, и его величина определяется минимумом полной энергии.

Предположение, что в цепочке имеется один электрон на атом, не принципиально. Если концентрация электронов n составляет долю α от концентрации атомов $1/a$, $n = \alpha/a$, то новая граница зоны Бриллюэна попадает на место раздела между занятыми и свободными состояниями при появлении в цепочке нового периода $\tilde{a} = 2a/\alpha$. Подобрать подходящий период можно при любой концентрации носителей. При этом можно даже обойтись без смещений атомов цепочки, создав для образования дополнительного периода волну зарядовой плотности в самой электронной системе. Поскольку при появлении дополнительного периода появляется щель в спектре между занятыми и свободными состояниями и металл превращается в изолятор, то отсюда следует утверждение, что одномерного металла при абсолютном нуле вообще не может быть. При повышении температуры выигрыш электронной энергии постепенно уменьшается из-за появления пустых состояний под щелью, где энергия состояний понижена по сравнению с невозмущенным спектром, и заполненных состояний над щелью, где энергия состояний повышена. Поэтому при некоторой критической температуре произойдет фазовый переход в металлическое состояние.

В рассуждении Пайелса одномерность существенна, потому что в системах с большей размерностью ферми-поверхность всегда имеет какую-то кривизну, а бриллюэновские поверхности разрыва энергии всегда плоские. Поэтому превратить трехмерный металл в изолятор за счет понижения симметрии не удастся. Однако понижение электронной энергии при появлении новых плоскостей разрыва часто оказывается существенным. Именно этим объясняется само существование полуметаллов Bi, Sb и As, а также существование металлических сплавов с очень сложными структурами, имеющими много атомов в элементарной ячейке. Кроме того, в анизотропных средах были обнаружены волны зарядовой и спиновой плотности, образование которых сопровождается фазовыми переходами. Все это вместе превратилось в отдельное направление в физике твердого тела, которое находится вне рамок данного курса. Поэтому о переходе Пайерлса мы здесь упоминаем лишь очень бегло, не коснувшись таких вопросов, как масштаб смещений атомов при $T = 0$, критическая температура перехода, при которой щель обратится в нуль, и т. п. (см., например, [2] или [3]).

А вообще мы о переходе Пайерлса упоминаем по двум причинам. Во-первых, на его примере удобно продемонстрировать принцип самовоздействия электронного газа на собственный энергетический спектр. Далее в этой главе мы познакомимся с тем, как этот принцип реализуется в разупорядоченной среде. Аналогия хорошо видна, например, при сравнении рис. 3.1 с рис. 3.7. Во-вторых, в начале гл. 5, где обсуждается понятие «чисто электронного фазового перехода», переход Пайерлса фигурирует в качестве отрицательного примера: хотя уменьшение электронной энергии является для него основным фактором, точка перехода определяется из энергетического баланса, в который входят и другие виды энергии, например, упругая.

3.2. Структура примесной зоны при слабом легировании

В свое время зонная теория успешно объяснила деление твердых тел с идеальной кристаллической решеткой на металлы и изоляторы. В металлах ферми-уровень находится внутри одной или нескольких перекрывающихся зон, т. е. в области разрешенных значений энергии. В изоляторе он попадает в область запрещенных значений энергии; при этом количество свободных носителей при некоторой температуре T определяется экспонентой, в показателе которой расстояние $|E_b - \mu|$ от ферми-уровня μ до ближайшего края разрешенных значений E_b сравнивается с T . Если $|E_b - \mu|$ имеет порядок атомных энергий (1 эВ, т. е. 10000 К), так что при комнатных температурах $T_r \approx 300$ К

$$|E_b - \mu| \gg T_r, \quad (3.1)$$

то проводимость при комнатной температуре неизмеримо мала и вещество является «истинным изолятором». Если же неравенство (3.1) не очень сильное, то имеется конечная проводимость, которая называется собственной. Собственная проводимость «вымерзает» при понижении температуры.

К вопросу о терминологии. Следуя традиции, сложившейся в физике полупроводников, в этой и следующей главах мы обозначаем ферми-уровень через μ ; в остальной части книги вместо этого употребляется термин ферми-энергия, обозначаемый через ε_F . И то и другое означает уровень энергии, который при абсолютном нуле температуры $T = 0$ отделяет занятые электронные состояния от свободных. В металлах состояния вблизи ферми-уровня делокализованы, в полупроводниках и изоляторах они могут быть локализованы или вообще отсутствовать. В последнем случае говорят, что ферми-уровень находится в запрещенной зоне.

Из находящихся в запрещенной зоне уровней энергии примесей E_0 наибольший интерес представляют те, которые находятся вблизи краев зон E_b . Они называются мелкими донорами или акцепторами. При их наличии ферми-уровень перемещается в окрестность E_0 , а количество свободных носителей и, соответственно, проводимость контролируются концентрацией примесей N и их энергией ионизации $|E_b - E_0|$ в сравнении с T . Соответствующая проводимость называется примесной.

Материалы, в которых имеется собственная или примесная проводимость, называются полупроводниками. Из сказанного понятно, что выделение этого класса веществ весьма условно и зависит, в частности, от температуры. При низких температурах это понятие «вымерзает» вместе с проводимостью. Однако название «полупроводники» прочно закрепилось за рядом веществ: Ge, Si, GaAs, InSb, InP и некоторыми другими.

Будем для определенности рассматривать полупроводник n -типа с шириной запрещенной зоны Δ . Пусть основная примесь в нем — мелкие водородоподобные доноры, т.е. доноры с энергией ионизации $E_0 \ll \Delta$ и с электронной волновой функцией, убывающей на больших расстояниях как $\exp(-r/a_B)$ (боровский радиус $a_B = \hbar(2m^*E_0)^{-1/2} = \hbar^2\kappa/m^*e^2$, где κ — диэлектрическая проницаемость, m^* — эффективная масса). Очистить и легировать полупроводник так, чтобы в нем были только доноры и совсем не было акцепторов, невозможно. Наличие какого-то количества акцепторов в полупроводнике n -типа имеет принципиальное значение. Поскольку энергия электрона на акцепторе ниже, чем на доноре, все N_A акцепторов захватят электроны с доноров и будут заряжены отрицательно, а из $N_D > N_A$ доноров N_A потеряют свои электроны и будут заряжены положительно. Поэтому там обязательно есть случайные электрические поля, созданные $2N_A$ хаотически

расположенными зарядами. Энергии всех доноров из-за этих полей изменятся:

$$\varepsilon_j = E_0 + \Delta\varepsilon_j, \quad \Delta\varepsilon_j = \frac{e^2}{\kappa} \left[\sum_a \frac{1}{|r_j - r_a|} - \sum'_{id} \frac{1}{|r_j - r_{id}|} \right]. \quad (3.2)$$

Первая сумма берется по всем акцепторам (переменный индекс a), а вторая — по всем ионизованным донорам (индекс id). Под индексом j понимается любой донор, штрих у второго знака суммирования означает, что когда индекс j относится к ионизованному донору, из суммы надо исключить член $j = id$ с нулем в знаменателе.

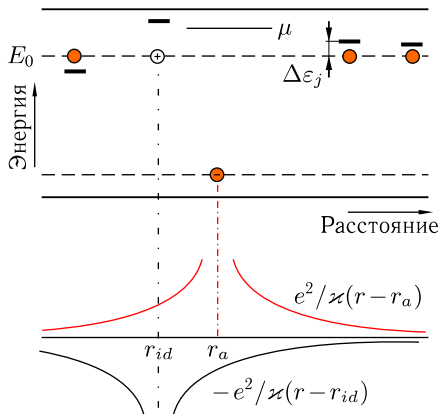


Рис. 3.2. Электрические поля заряженных центров. Вверху на энергетической схеме кружками отмечены изначальные значения энергии доноров и акцептора, горизонтальными черточками — значения энергии доноров, скорректированные кулоновским потенциалом заряженных примесей. Этот потенциал показан внизу

Наличие сдвигов $\Delta\varepsilon_i$ означает, что даже в слабо легированном полупроводнике ($Na_B^3 \ll 1$) из-за хаотических электрических полей N_D -кратно вырожденный уровень донора E_0 расплывается в зону. Она называется примесной зоной. Ферми-уровень должен разделять заполненные уровни нейтральных и пустые уровни ионизованных доноров, т. е. находиться внутри этой примесной зоны (см. рис. 3.2).

Заметьте. Это превращение уровня в примесную зону при $Na_B^3 \ll 1$ не связано с перекрытием волновых функций; оно имеет чисто классическую причину. Перекрытие, а при какой-то степени перекрытия и делокализация, возникают при более сильном легировании. Пока этого не произошло, все состояния в такой примесной зоне локализованы, а сам термин «зона» означает интервал энергий с отличной от нуля плотностью состояний.

Слабая компенсация. Рассмотрим случай, когда акцепторов мало, так что коэффициент компенсации $K = N_A/N_D \ll 1$. Поскольку разноименные заряды притягиваются, в равновесии ионизованы будут те доноры, которые находятся близко к акцепторам. Неравенство $N_A \ll N_D$ означает, что среднее расстояние между акцепторами много больше, чем среднее расстояние между донорами. Поэтому каждый ионизованный донор можно приписать к определенному акцептору и свести задачу к рассмотрению возможных локальных конфигураций заряженных доноров вокруг одного заряженного акцептора.

Если близко от какого-то акцептора расположены два донора, то они могут быть ионизованы оба. Пусть акцептор, который у нас всегда заряжен отрицательно, расположен посередине между двумя донорами на расстоянии r от каждого (рис. 3.3, а) и пусть один из доноров ионизован, т.е. заряжен положительно. Приход положительного заряда из бесконечности на второй донор даст выигрыш в энергии $e^2/\kappa r$ благодаря притяжению к акцептору, а проигрыш из-за отталкивания от первого донора $e^2/2\kappa r$ будет в два раза меньше. Следовательно такая конфигурация возможна и один акцептор может удерживать около себя два ионизованных донора.

Удерживать около себя три ионизованных донора один акцептор не может. В самой благоприятной конфигурации доноры расположены в углах равностороннего треугольника (со стороной r), а акцептор в центре (рис. 3.3, б). Если два донора уже ионизованы, то ионизация третьего, описываемая как приход на него положительного заряда из бесконечности, даст проигрыш в энергии:

$$\frac{e^2}{\kappa r} (\sqrt{3} - 2) < 0.$$

Сказанное позволяет разбить акцепторы на три группы. N_0 акцепторов вообще не имеют около себя ионизованных доноров, N_1 имеют по одному, а N_2 — по два ионизованных донора. Из условия электронейтральности $N_0 + N_1 + N_2 = N_A$ следует уравнение для уровня Ферми:

$$N_0(\mu) = N_2(\mu). \quad (3.3)$$

Напишем выражение для $N_0(\mu)$. Пусть нуль на шкале энергий соответствует энергии изолированного донора. Как видно из рис. 3.2, для того чтобы акцептор принадлежал к группе N_0 , около него не должно быть доноров в радиусе $r_\mu = e^2/\kappa\mu$. Поэтому N_0 есть умноженная

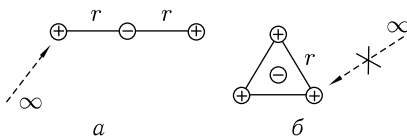


Рис. 3.3. Реализуемые (а) и нереализуемые (б) конфигурации заряженных центров

на N_A вероятность того, что донора нет в объеме $(4\pi/3)r_\mu^3$:

$$N_0(\mu) = N_A \exp\left(-\frac{4\pi}{3} r_\mu^3 N_D\right) = N_A \exp\left[-\frac{4\pi}{3} \left(\frac{\varepsilon_D}{\mu}\right)^3\right], \quad (3.4)$$

где $\varepsilon_D = (e^2/\kappa)N_D^{1/3}$ — энергия кулоновского взаимодействия на среднем расстоянии между донорами $N_D^{-1/3}$.

Вычисление $N_2(\mu)$ несколько сложнее. Вероятность того, что два ближайших к акцептору донора расположены на расстояниях $r_1 = |\mathbf{r}_1|$ и $r_2 = |\mathbf{r}_2| \geq r_1$ в элементах объема $d\mathbf{r}_1$ и $d\mathbf{r}_2$ равна

$$\begin{aligned} N_D^2 \exp\left(-\frac{4\pi}{3} r_1^3 N_D\right) \exp\left(-\frac{4\pi}{3} (r_2^3 - r_1^3) N_D\right) d\mathbf{r}_1 d\mathbf{r}_2 = \\ = N_D^2 \exp\left(-\frac{4\pi}{3} r_2^3 N_D\right) d\mathbf{r}_1 d\mathbf{r}_2. \end{aligned} \quad (3.5)$$

Чтобы оба донора были ионизованы, уровни электронов на обоих должны быть выше μ :

$$\varepsilon_1 = \frac{e^2}{\kappa|\mathbf{r}_1|} - \frac{e^2}{\kappa|\mathbf{r}_1 - \mathbf{r}_2|} > \mu, \quad \varepsilon_2 = \frac{e^2}{\kappa|\mathbf{r}_2|} - \frac{e^2}{\kappa|\mathbf{r}_1 - \mathbf{r}_2|} > \mu. \quad (3.6)$$

Тогда

$$\begin{aligned} N_2(\mu) = N_A N_D^2 \int d\mathbf{r}_1 \int_{r_2 > r_1} d\mathbf{r}_2 \exp\left(-\frac{4\pi}{3} r_2^3 N_D\right) \Theta(\varepsilon_1 - \mu) \Theta(\varepsilon_2 - \mu), \\ \Theta(x) = \begin{cases} 1 & \text{при } x > 0, \\ 0 & \text{при } x < 0. \end{cases} \end{aligned} \quad (3.7)$$

Вычисленная по формуле (3.7) величина $N_2(\mu)$, вообще говоря, несколько занижена: может оказаться, что два донора, ближайших к акцептору, не могут быть одновременно ионизованы, так как они находятся слишком близко друг к другу. Такой акцептор в (3.7) учтен не будет. Однако по другую сторону от этого акцептора может быть расположен еще один донор на расстоянии $r_3 > r_2$, который тем не менее может составить ионизованную пару с одним из ближайших доноров. Однако соответствующая поправка очень мала, менее одного процента. Подставив в уравнение (3.3) результаты вычислений (3.7) с поправкой и (3.4), получим величины $N_0 = N_2$ и N_1 и положение ферми-уровня:

$$N_0 = N_2 = 0,0013 N_A, \quad N_1 = 0,974 N_A, \quad (3.8)$$

$$\mu \approx 0,99 \varepsilon_D. \quad (3.9)$$

График плотности состояний схематически представлен слева на рис. 3.4.

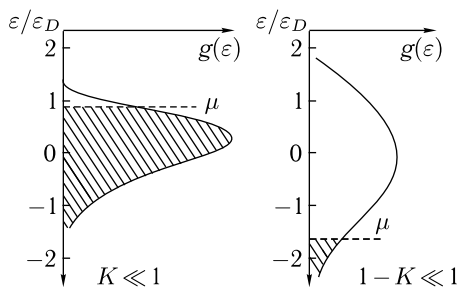


Рис. 3.4. Плотность состояний в примесной зоне при слабой ($K \ll 1$) и при сильной ($1 - K \ll 1$) компенсациях. Заштрихованы заполненные состояния, энергия отсчитывается от уровня E_0

Заметьте. Площадь под кривой $g(\varepsilon)$ равна N_D , а эффективная ширина порядка ε_D . Смысл коэффициента 0,99 в соотношении (3.9) чисто символический: показать, что он возник в результате расчета и есть нечто большее, нежели просто «порядка единицы». В то же время этот расчет ничего не говорит о степени симметричности кривой $g(\varepsilon)$ относительно E_0 .

Зависимость $\mu - E_0$ от N_D можно экспериментально проверить благодаря тому, что эта разность есть энергия активации при прыжковой проводимости по ближайшим центрам (относительно прыжковой проводимости см. гл. 4). На рис. 3.5 приведены данные двух разных экспериментов на п-Ge. Видно, что по крайней мере при малых N_D эксперимент подтверждает зависимость (3.9). При больших N_D , вероятно,

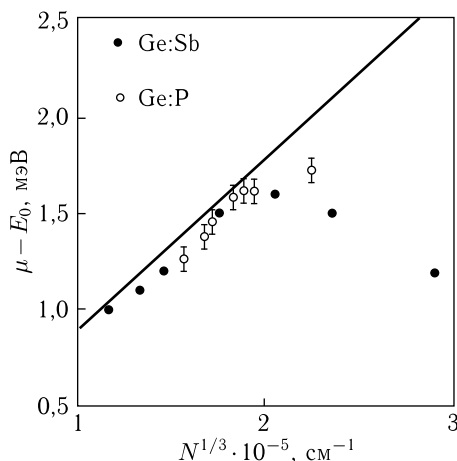


Рис. 3.5. Энергия активации при прыжковой проводимости по ближайшим центрам в зависимости от концентрации примесей, Ge:P — [4]; Ge:Sb — [5].

Прямой показана теоретическая зависимость (3.9)

становится существенным перекрытие волновых функций электронов на центрах, которое не учитывалось описанным выше классическим расчетом.

Заметьте. И плотность состояний, и положение ферми-уровня сформировались в том потенциальном рельефе, который был создан самими электронами, в принципе имевшими возможность распределиться по донорам $C_{N_D}^{N_A}$ способами. Приведенное выше решение было получено в пределе $K \ll 1$; оно описывает функцию $g(\varepsilon)$ на масштабе ε_D . По мере роста K растет и число $C_{N_D}^{N_A}$ способов размещения электронов на донорах, что дает основания ожидать усложнение функции $g(\varepsilon)$. И действительно, как показано в параграфе 3.3, у функции $g(\varepsilon)$ возникает дополнительная структура с меньшим масштабом в окрестности μ .

Сильная компенсация. Рассмотрим другой предельный случай

$$1 - K \ll 1, \quad n = N_D - N_A \ll N_D. \quad (3.10)$$

К вопросу о терминологии. Буквой n у нас везде обозначается количество носителей, а если электроны локализованы, то количество «потенциальных носителей». В соответствии с этим здесь $n = N_D - N_A$ определяет количество электронов, оставшихся в полосе донорных состояний; в противоположность им N_A электронов и при слабой, и при сильной компенсации ушли на акцепторы и вообще как бы вышли из игры. Под игрой (под будущей игрой) здесь понимаются прыжковая проводимость при низкой но ненулевой температуре, переход металл–изолятор при увеличении n , и т. д.

В условиях (3.10) большая часть доноров ионизована. Нейтральным останется лишь тот донор, который расположен слишком близко от другого ионизованного донора: в парах близко расположенных доноров один остается нейтральным, поскольку его энергия понижена на $\varepsilon = e^2/\kappa r$ в поле положительно заряженного соседа (здесь r — расстояние между донорами в паре).

Вероятность иметь близкого соседа на расстоянии от r до $r + dr$ есть

$$\nu_r dr = N_D \cdot 4\pi r^2 dr, \quad (3.11)$$

вероятность понизить энергию на ε за счет соседа равна

$$\nu_\varepsilon d\varepsilon = N_D \cdot 4\pi r^2 \frac{dr}{d\varepsilon} d\varepsilon = -N_D \cdot 4\pi \frac{(e^2/\kappa)^3}{\varepsilon^4} d\varepsilon. \quad (3.12)$$

Умножив на N_D и разделив на 2, чтобы не учитывать один донор дважды, получим плотность состояний для пар

$$g_2(\varepsilon) = 2\pi N_D \frac{\varepsilon_D^3}{\varepsilon^4}, \quad \varepsilon_D = \frac{e^2 N_D^{1/3}}{\kappa}. \quad (3.13)$$

Уровень Ферми находится из условия

$$\int_{-\infty}^{\mu} g_2(\varepsilon) d\varepsilon = n. \quad (3.14)$$

Отсюда

$$\mu = -\left(\frac{2\pi}{3}\right)^{1/3} \varepsilon_D \left(\frac{N_D}{n}\right)^{1/3} = -\left(\frac{2\pi}{3}\right)^{1/3} \frac{\varepsilon_D}{(1-K)^{1/3}}. \quad (3.15)$$

(см. схематический график в правой части рис. 3.4).

Формула (3.15) содержит, помимо N_D , еще один переменный параметр, а именно $(1-K)$. Зависимость от него тоже может быть проверена по измерениям энергии активации прыжковой проводимости по ближайшим центрам. Сравнение приведено на рис. 3.6.

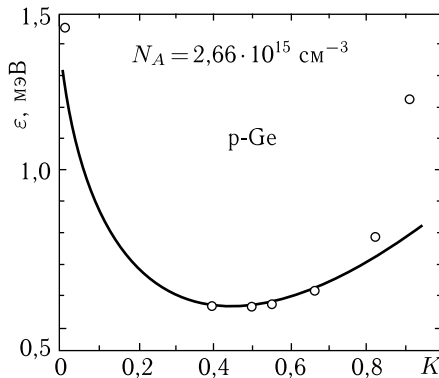


Рис. 3.6. Энергия активации при прыжковой проводимости по ближайшим центрам в зависимости от коэффициента компенсации [6]. В правой части, при $K > 0,5$, теоретическая кривая выходит на зависимость (3.15)

3.3. Кулоновская щель

В пределах слабой $K \ll 1$ и сильной $1-K \ll 1$ компенсации получилась колоколообразная плотность состояний $g(\varepsilon)$ с максимумом вблизи E_0 и с ферми-уровнем на одном из крыльев распределения (рис. 3.4). Можно было бы ожидать, что при постепенном росте K ферми-уровень плавно смещается с одного края мало меняющейся функции $g(\varepsilon)$ на другой, приходясь на ее максимум при $K \approx 0,5$, когда электронов в 2 раза меньше, чем центров в примесной зоне.

То, что это не так, видно из приведенного на рис. 3.7 результата компьютерного моделирования для $K = 0,5$. Плотность состояний на

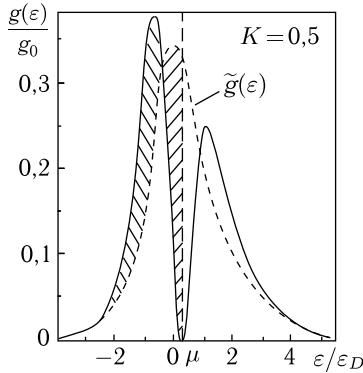


Рис. 3.7. Плотность состояний в примесной зоне при коэффициентах компенсации $K = 0,5$ по результатам компьютерного моделирования [7]. Заштрихованные площади равны (ср. с рис. 3.1).

этом графике отложена в единицах g_0 , а энергия в единицах ε_D :

$$g_0 = \frac{N_D}{\varepsilon_D}, \quad \varepsilon_D = \frac{e^2}{\kappa} N_D^{1/3}. \quad (3.16)$$

Ферми-уровень μ действительно оказался близко к $E_0 = 0$, но при значении $\varepsilon = \mu$ у функции $g(\varepsilon)$ появился минимум $g(\mu) = 0$.

Этот результат требуется осмыслить и объяснить.

Предпосылкой для такого результата является большое количество $C_{N_D}^{N_A}$ вариантов распределения электронов по центрам: есть из чего выбирать.

Нетрудно объяснить, *почему* это распределение оказалось предпочтительным. Сравним его с «ожидавшимся», с показанной на рис. 3.7 пунктиром кривой $\tilde{g}(\varepsilon)$ с теми же крыльями и той же площади

$$\int_{-\infty}^{\infty} \tilde{g}(\varepsilon) d\varepsilon = \int_{-\infty}^{\infty} g(\varepsilon) d\varepsilon = N_D, \quad (3.17)$$

но без минимума. Нетрудно видеть, что средняя энергия $\bar{\varepsilon}_-$ состояний ниже уровня Ферми уменьшилась:

$$\bar{\varepsilon}_- = \frac{1}{N_D - N_A} \int_{-\infty}^{\mu} \varepsilon g(\varepsilon) d\varepsilon < \frac{1}{N_D - N_A} \int_{-\infty}^{\mu} \varepsilon \tilde{g}(\varepsilon) d\varepsilon, \quad (3.18)$$

так как центр тяжести этой части распределения сместился в область меньших энергий. Поскольку при низких температурах занятые состояния находятся преимущественно в этой части спектра, полная энергия

электронного газа

$$\hat{E} = \int_{-\infty}^{\infty} \varepsilon g(\varepsilon) f\left(\frac{\varepsilon}{T}\right) d\varepsilon, \quad f(x) = (\exp x + 1)^{-1}, \quad (3.19)$$

уменьшится, несмотря на то, что средняя энергия $\bar{\varepsilon}_+$ состояний в части спектра $\varepsilon > \mu$ увеличилась. Эти рассуждения похожи на те, которые объясняют наличие перехода Пайерлса в одномерной системе. Сравнивая изменения плотности состояний от $\tilde{g}(\varepsilon)$ к $g(\varepsilon)$ из-за неустойчивости Пайерлса (рис. 3.1) и при образовании кулоновской щели (рис. 3.7), нетрудно видеть, что уравнения (3.17)–(3.19) применимы в обоих случаях.

Заметьте. Из этой аналогии видно, что температура может изменять плотность состояний $g(\varepsilon)$ и кулоновскую щель.

Для того чтобы показать, как и насколько изменяется функция $g(\varepsilon)$ вблизи μ , воспользуемся следующим рассуждением. Пусть в равновесии при $T = 0$ два состояния из интервала энергий $(\mu - \varepsilon/2, \mu + \varepsilon/2)$, одно заполненное: $\varepsilon_i < \mu$, а другое пустое: $\varepsilon_j > \mu$, находятся на расстоянии r_{ij} . Унесем электрон с центра i на бесконечность. Тогда энергия уровня j понизится до $\varepsilon_j - e^2/\kappa r_{ij}$. Но она все равно должна остаться больше ε_i , иначе изначально более выгодной была бы конфигурация «центр j заполнен, центр i пустой». Из

$$\varepsilon_j - \varepsilon_i - \frac{e^2}{\kappa r_{ij}} > 0, \quad \varepsilon_j - \varepsilon_i < \varepsilon \quad (3.20)$$

следуют зависящие от ε ограничения снизу для r_{ij} и сверху для N :

$$\varepsilon > \frac{e^2}{\kappa r_{ij}}, \quad N(\varepsilon) = r_{ij}^{-3} < \left(\frac{\kappa \varepsilon}{e^2}\right)^3. \quad (3.21)$$

Поскольку энергия ε принимает здесь произвольные, сколь угодно малые значения, заменяя знак неравенства на знак равенства и дифференцируя, получаем

$$g(\varepsilon) = \frac{\partial N}{\partial \varepsilon} \propto \frac{\kappa^3}{e^6} (\varepsilon - \mu)^2. \quad (3.22)$$

Использование знака равенства означает, что мы выбрали самую малую степень $(\varepsilon - \mu)$ из всех, удовлетворяющих соотношениям (3.21).

Такое поведение функции $g(\varepsilon)$ вблизи μ называется мягкой кулоновской щелью. «Мягкой» потому, что $g(\varepsilon)$ обращается в нуль только в одной точке, а «кулоновской» потому, что она возникла в результате кулоновского взаимодействия.

Формулы (3.21) и (3.22) написаны для трехмерной среды. В двумерном случае из аналогичных рассуждений следует

$$N(\varepsilon) = r_{ij}^{-2} < \left(\frac{\kappa \varepsilon}{e^2}\right)^2, \quad g(\varepsilon) = \frac{\partial N}{\partial \varepsilon} \propto \frac{\kappa^2}{e^4} |\varepsilon - \mu|. \quad (3.23)$$

Во всех рассуждениях по поводу кулоновской щели коэффициент компенсации K в явном виде не фигурировал. Следует полагать, что функция $g(\varepsilon)$ обращается в нуль на ферми-уровне не только при $K \approx 0,5$, но и при значениях K , расположенных ближе к краям интервала $(0, 1)$. Это подтверждается результатами компьютерного моделирования, представленными на рис. 3.8 в том же самом масштабе (3.16), что и на рис. 3.7.

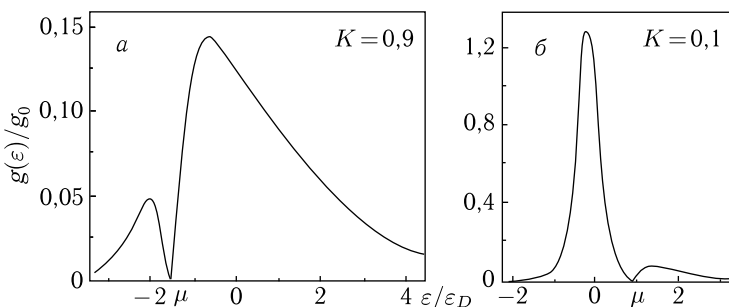


Рис. 3.8. Плотность состояний в примесной зоне при коэффициентах компенсации K , равных 0,1 и 0,9, по результатам компьютерного моделирования [7]

Заметьте. Распределение при $K = 0,9$ шире, чем при $K = 0,1$. Этого следовало ожидать из сравнения формул (3.9) и (3.15): в (3.15) имеется дополнительный численный множитель около 1,3 и множитель $(1 - K)^{1/3}$ в знаменателе. Хотя формулы (3.9) и (3.15) не содержат кулоновской щели, ширину кривых и положение ферми-уровня они отражают правильно.

Кулоновская щель непосредственно влияет на температурную зависимость прыжковой проводимости электронов в окрестности ферми-уровня (прыжки с переменной длиной). Это позволяет экспериментально изучать кулоновскую щель. Мы подробно обсуждаем этот вопрос в гл. 4 о прыжковой проводимости. Там же приведены и экспериментальные кривые, которые, в частности, дают пищу для размышлений о том, в какой мере кулоновская щель — универсальное явление, и существуют ли изоляторы с постоянной, отличной от нуля плотностью состояний в окрестности ферми-уровня.

Прыжковая проводимость — это косвенная демонстрация существования кулоновской щели. Существует и более прямой метод — измерение вольт-амперных характеристик туннельных контактов. Из таких непосредственно измеренных характеристик можно извлечь функцию плотности состояний в окрестности ферми-уровня. Обычно свойства исследуемого материала постепенно меняют, начиная с металлического состояния и кончая изолятором. В ходе такой трансформации на ферми-уровне появляется минимум в плотности состояний, который расширяется и углубляется, доходя до $g(\mu) = 0$. Обычно параллельно

наблюдают радикальные изменения характера температурной зависимости проводимости. В интерпретацию таких экспериментов вовлекается не только кулоновская щель, но и особенности межэлектронного взаимодействия в грязных металлах, переходы металл-изолятор, прыжковая проводимость и т. п. (см. Приложение Б). Во всех этих вопросах кулоновское взаимодействие играет существенную роль.

Почувствительно проследить развитие понятия кулоновской щели. Оно появилось как реакция на экспериментальные исследования прыжковой проводимости, которые обнаружили температурные зависимости проводимости, не соответствующие ни активационному закону Аррениуса (4.13), ни закону Мотта (4.19) при переменной длине прыжка и постоянной плотности состояний в окрестности ферми-уровня. Введение понятия кулоновской щели позволило последовательно и достаточно полно описать имеющиеся экспериментальные данные по прыжковой проводимости (см. гл. 4). Более того, предсказанная формулами (3.22) и (3.23) форма кулоновской щели была экспериментально подтверждена в туннельных экспериментах.

На рис. 3.9 и рис. 3.10 представлены туннельные характеристики, полученные на контактах, в которых один электрод — металл, а другой — изолятор. Как показано в Приложении Б, величина dI/dV , отложенная по оси ординат на этих характеристиках, пропорциональна плотности состояний в изоляторе, а приложенное к контакту напряжение определяет отсчитанную от ферми-уровня энергию, к которой относится эта плотность состояний. В обоих изоляторах обнаруживается кулоновская щель, причем в объемном Si:B щель параболическая, в соответствии с формулой (3.22), а в сверхтонких пленках Be она меняется линейно с энергией, как это и предсказывается формулой

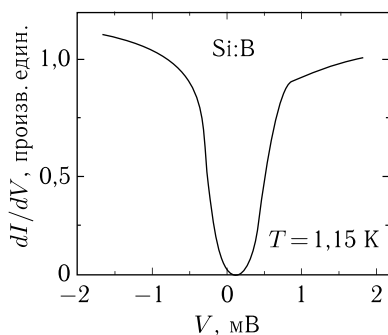


Рис. 3.9. Дифференциальная проводимость туннельного контакта с Si:B, находящимся в состоянии изолятора из-за малого количества носителей [8]. Полную серию кривых см. в Приложении Б на рис. 5

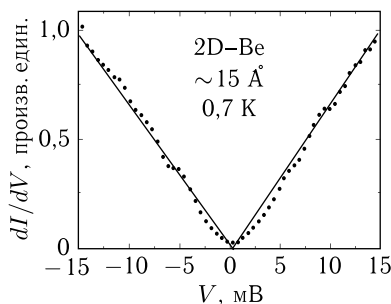


Рис. 3.10. Дифференциальная проводимость туннельного контакта с ультратонкой высокоомной пленкой Be при $T = 0,7$ К [9]. Полную серию кривых см. в Приложении Б на рис. 6

(3.23). Более подробно измерения на обоих веществах обсуждаются в Приложении Б.

Вместе с тем на теоретическом уровне с самого начала оставалось некоторое чувство неудовлетворенности. Рассуждения (3.20)–(3.22) эквивалентны учету только парного межэлектронного взаимодействия, поскольку при перемещении электрона с центра i на центр j расположение остальных электронов предполагалось неизменным. Между тем это перемещение могло бы стимулировать цепочку перемещений других электронов. Тогда следовало бы сравнивать между собой энергии конфигураций, отличающихся числами заполнений не двух, а большего количества центров. Это должно было бы существенно повлиять на электронный спектр. Поэтому решающее слово в вопросе о форме кулоновской щели за экспериментом. Мы вернемся к этому вопросу в следующей главе.

3.4. Контрольные вопросы и задачи

1. Какова вероятность, что в материале с концентрацией доноров N_D не окажется ни одного донора в объеме V ?

$$\exp(-N_D V).$$

2. Изменяются ли число реализуемых конфигураций (рис. 3.3) и формулы (3.3)–(3.9), описывающие примесную зону, при переходе от 3-к 2-мерному случаю? Пленку какой толщины следует считать двумерной?

Число различных конфигураций не изменится, а конкретные выражения будут другие.

$$b \ll N_D^{1/2}.$$

3. Оцените плотность состояний в примесной зоне в точках $\varepsilon/\varepsilon_D = 0$ и $\varepsilon/\varepsilon_D = \mu$ при слабой и сильной компенсации.

$$N_D/\varepsilon_D \text{ и } N_A/\varepsilon_D.$$

4. Пусть концентрация доноров N_D больше концентрации акцепторов N_A , $N_D > N_A$. Предположим, что около каждого акцептора имеется один донор и что полем получившихся диполей на других донорах можно пренебречь. Как будет выглядеть спектр примесной зоны доноров при таком предположении?

$$g(\varepsilon) = (N_D - N_A)\delta(0) + \frac{4\pi\varepsilon_D^3}{\varepsilon^4} N_A \exp\left(-\frac{4\pi N_D \varepsilon^6}{3\kappa^3 \varepsilon^3}\right).$$

Первый член — нейтральные доноры, второй — ионизованные в составе диполей.

5. Пусть в нашей системе имеются доноры двух сортов, N_{D1} и N_{D2} , с несколькими разными энергиями ионизации, $N_D = N_{D1} + N_{D2} \gg N_A$. Нарисуйте энергетический спектр.

6. Возможна ли зависимость кулоновской щели от температуры? Да.

Список литературы

1. Шкловский Б. И., Эфрос А. Л. Электронные свойства легированных полупроводников. — Наука, 1979.
2. Пайерлс Р. Квантовая теория твердых тел. — ИЛ, 1956.
3. Grüner G. Density Waves. — Perseus Books, 2000.
4. Шкловский Б. И., Шлимак И. С. // ФТП **6**, 129 (1972).
5. Fritzsche H. // J. Phys. Chem. Sol. **6**, 69 (1958).
6. Mott N. F., Twose W. D. // Adv. Phys. **10**, 707 (1963) и УФН **79**, 691 (1963) по данным H. Fritzsche и M. Guevas
7. Efros A. L., Lien N. V., Shklovskii B. I. // J. Phys. C **12**, 1023 (1979).
8. Massey J. G., Lee M. // Phys. Rev. Lett. **77**, 3399 (1996).
9. Butko V. Yu., DiTusa J. F., Adams P. W. // Phys. Rev. Lett. **84**, 1543 (2000).

Глава 4

ПРЫЖКОВАЯ ПРОВОДИМОСТЬ ¹⁾

4.1. Локализованные состояния и переходы между ними

В этой главе мы будем говорить только о локализованных состояниях. Определим их как состояния, волновые функции которых сосредоточены в основном в ограниченной области пространства и экспоненциально спадают с расстоянием вне этой области:

$$\psi \rightarrow f(r)e^{-r/\xi} \quad \text{при} \quad r \rightarrow \infty.$$

Коэффициент ξ в показателе экспоненты называется радиусом локализации или локализационной длиной. Например, в центрально-симметричной трехмерной потенциальной яме

$$U(r) = \begin{cases} 0, & r > a, \\ -U_0, & r \leq a \end{cases}$$

волновые функции на бесконечности убывают как

$$\psi \propto \frac{1}{r} e^{-r/\xi}, \quad \xi = \frac{\hbar}{\sqrt{2m|E|}}, \quad (4.1)$$

где E — положение уровня, отсчитанное от верхнего края ямы.

Волновые функции электрона на уровнях одномерной прямоугольной ямы затухают по закону

$$\psi(x) \propto \exp(-x/\xi), \quad (4.2)$$

т. е. $f(x) = \text{const.}$ Радиус локализации определяется тем же выражением (4.1), что и в трехмерном случае. В частности, в мелкой яме шириной a_w и глубиной $U_0 \ll \hbar^2/(ma_w^2)$, в которой имеется всего один уровень, волновая функция затухает на длине $\xi = \hbar^2/(ma_w U_0)$.

¹⁾ Несколько иное изложение тех же вопросов можно найти в книге [1]; прыжковой проводимости посвящена большая коллективная монография [2].

Обратите внимание в последнем примере на различие между размером ямы a_w и длиной ξ , на которой затухает волновая функция электрона.

Длина локализации в притягивающем кулоновском потенциале $U = -e^2/r$, создаваемом единичным зарядом e , называется боровским радиусом $a_B = \hbar^2/(me^2)$. Уровни энергии E_n и асимптотика волновых функций электрона в атоме водорода ψ_n зависят от главного квантового числа n и равны

$$E_n = \frac{E_1}{n^2}, \quad \psi_n(r) \rightarrow C(n)r^{n-1} \exp\left(-\frac{r}{na_B}\right) \quad \text{при } r \rightarrow \infty \quad (4.3)$$

$$(n = 1, 2, 3, \dots),$$

где $C(n)$ — это константа. Из выражения (4.3) видно, что локализационная длина электрона в атоме водорода зависит от главного квантового числа, $\xi_n = na_B$. В основном состоянии $\xi_1 = a_B$.

При описании примесных центров в твердых телах, в частности, мелких доноров и акцепторов, о которых шла речь в предыдущей главе, обычно используется модель водородоподобных центров с боровским радиусом

$$a_B = \frac{\kappa \hbar^2}{m^* e^2}, \quad (4.4)$$

в выражение для которого дополнительно входит диэлектрическая проницаемость κ , а в знаменателе стоит вместо обычной эффективная масса электрона m^* .

При конечных температурах локализованные состояния вносят вклад в процессы переноса. Перенос по локализованным состояниям возникает в результате прыжков носителей с одного локализационного центра на другой и поэтому называется прыжковой проводимостью. Важным постулатом всей концепции прыжковой проводимости является предположение, что все центры практически имеют различные энергии: два центра с одинаковой энергией находятся на бесконечном расстоянии друг от друга. При описании прыжковой проводимости будем исходить из модели примесной зоны при слабом легировании, считая, что примесные центры имеют координаты r_i .

Пусть концентрация примесей N мала, так что $Na_B^3 \ll 1$ и волновые функции электронов локализованы. Но благодаря экспоненциальным хвостам волновых функций перекрытие соседних примесных центров все же имеется, и в меру этого перекрытия существует конечная вероятность перехода (прыжка) электрона с одного центра на другой:

$$\frac{1}{\tau_{ij}} \propto F(\varphi_{ij}, f_i, f_j) \int |M_q|^2 \delta(\hbar q s - \Delta_{ij}) d^3 q \propto$$

$$\propto F(\varphi_{ij}, f_i, f_j) \left| \int \psi_j^* e^{iqr} \psi_i d^3 r \right|^2 \quad (4.5)$$

(s здесь скорость звука). Множитель $\exp(iqr)$ под интегралом в (4.5) — это волновая функция фонона, а дельта-функция обеспечивает выполнение закона сохранения энергии и отбирает фононы, которые нужно испустить или поглотить, чтобы компенсировать разницу энергий начального и конечного состояний $\Delta_{ij} = \varepsilon_i - \varepsilon_j$. В первом приближении квадрат интеграла сводится к множителю $\exp(-2r_{ij}/a_B)$. Функция F аккумулирует в себе все статистические факторы, которые обуславливают переход: согласно принципу Паули переход возможен только на свободный уровень ε_j , поглотить фонон можно только, если он имеется в системе, и т. д. Аргументами функции F являются фермиевская

$$f_i = \frac{1}{\exp \frac{\varepsilon_i - \mu}{T} + 1}$$

и бозевская

$$\varphi_{ij} = \frac{1}{\exp \frac{\Delta_{ij}}{T} - 1}$$

функции распределения. В результате преобразований и пренебрежения второстепенными членами функция F обычно сводится к множителю $\exp(-\varepsilon_{ij}/T)$, где ε_{ij} — некоторая характерная энергия (не обязательно равная $\Delta_{ij} = \varepsilon_i - \varepsilon_j$). Этот множитель можно написать, исходя из физических соображений.

Таким образом, вероятность переходов $1/\tau_{ij}$ между центрами i и j пропорциональна произведению двух экспонент. Соединим каждую пару примесных центров воображаемым сопротивлением R_{ij} , обратно пропорциональным вероятности переходов:

$$R_{ij} = R_0 e^{u_{ij}}, \quad u_{ij} = \frac{2r_{ij}}{a_B} + \frac{\varepsilon_{ij}}{T}. \quad (4.6)$$

Получается сетка случайных сопротивлений, моделирующая изолятор. Она называется сеткой Абрахамса–Миллера. В равновесии электронные переходы между узлами сетки происходят с одинаковой вероятностью в обе стороны (принцип детального равновесия). Во внешнем электрическом поле появляется направленный поток электронов, т. е. конечная проводимость. Чтобы ее вычислить, нужно решить задачу о проводимости такой сетки.

На первый взгляд сетка Абрахамса–Миллера выглядит устрашающе, поскольку формально изначально каждый узел соединен со всеми остальными. Однако сопротивления между узлами, далеко расположенными друг от друга, экспоненциально велики; их можно без ущерба выкинуть, поскольку они заведомо зашунтированы существенно меньшими сопротивлениями. На этом основан общий принцип решения задачи. Из сетки последовательно убираются самые большие сопротивления до тех пор, пока сохраняется ее односвязность. Сопротивление сетки зависит от того, какие самые большие сопротивления пришлось оставить для сохранения связности. Это типичная задача теории про-

текания. Поэтому для понимания материала этой главы нужно ознакомиться с Приложением А (см., в частности, параграф А.4). Ниже мы будем обсуждать разные варианты прыжковой проводимости. В каждом варианте выделяется система существенных узлов, вычисляются значения R_{ij} и решается соответствующая перколяционная задача.

Заметьте. Удельное сопротивление ρ регулярной ортогональной сетки с размером ячейки a и сопротивлением связи R зависит от размерности сетки d :

$$\begin{aligned} d = 1 : \quad \rho &= R/a, \\ d = 2 : \quad \rho &= R, \\ d = 3 : \quad \rho &= Ra. \end{aligned} \tag{4.7}$$

Эти же соотношения справедливы и для произвольной сетки, но под a и R подразумеваются средние величины.

4.2. Прыжки на ближайшие центры

Наиболее простой вид прыжковой проводимости реализуется при переходах между ближайшими соседними центрами. Плотность состояний в примесной зоне доноров при малой компенсации имеет вид кривой с максимумом при энергии порядка энергии ионизации изолированного донора E_D . Наиболее вероятно, что исходный и конечный пункты прыжка, если они ближайшие соседи, имеют энергии вблизи максимума плотности состояний. Для того чтобы прыжок реализовался, конечный пункт должен быть свободен. Вероятность такого события зависит от расстояния до ферми-уровня μ и пропорциональна

$$\exp\left(-\frac{|\mu - E_D|}{T}\right), \tag{4.8}$$

и это самый малый множитель, входящий в функцию F в уравнении (4.5). Множитель $\exp(-\Delta_{ij}/T)$, определяющий количество участвующих в переходах фононов, гораздо больше, поскольку $\Delta_{ij} \ll |\mu - E_D|$. Поэтому величина u_{ij} , входящая в выражение (4.6), для всех прыжков определяется расстоянием от максимума плотности состояний до ферми-уровня: $\varepsilon_{ij} = |\mu - E_D|$.

Заметьте. Это как раз тот самый случай, когда $\varepsilon_{ij} \neq \Delta_{ij}$.

Поскольку фактор (4.8) входит во все u_{ij} , он не вносит вклада в разброс R_{ij} и может быть «вынесен за скобки», т. е. не учитываться при анализе свойств сетки:

$$R_{ij} = R_0 \exp \frac{2r_{ij}}{a_B}$$

(чем ближе друг к другу расположены узлы, тем меньше сопротивление между ними в модельной сетке).

Задача свелась к определению перколяционного радиуса r_c в системе случайно расположенных узлов, имеющих концентрацию N . Из теории перколяции известно (см. формулу (А.10) в Приложении А), что

$$\frac{4\pi}{3} r_c^3 N = B_c = 2,7, \quad (4.9)$$

откуда

$$r_c = 0,865 N^{-1/3} \quad \text{и} \quad \rho = \rho_0 \exp \frac{1,73}{N^{1/3} a_B}. \quad (4.10)$$

Обратимся теперь к эксперименту. На рис. 4.1 приведены температурные зависимости сопротивления германия при различной концентрации примесей. У этого эксперимента две особенности. Во-первых огромные диапазоны изменения измеренных сопротивлений (12 порядков!) и концентраций примесей (больше 3-х порядков). Во-вторых,

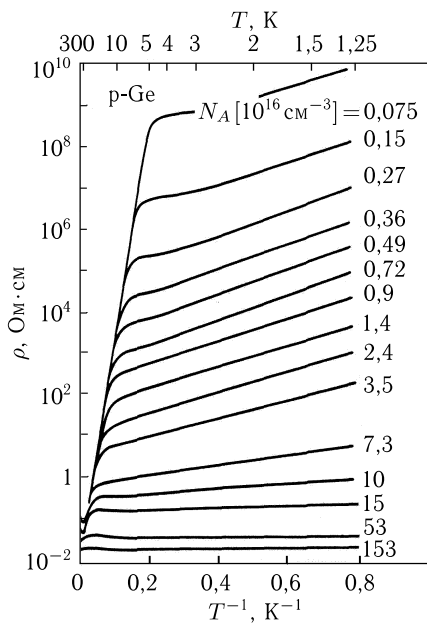


Рис. 4.1. График сопротивления образцов p-Ge с различной концентрацией примесей [3]

замечательный способ введения примесей, обеспечивающий отсутствие корреляций в их расположении и строгое поддержание коэффициента компенсации. Чистый Ge облучается в реакторе потоком нейтронов. В результате ядерных реакций с нейтронами из ядер одних изотопов Ge получаются ядра Ga, т.е. мелкие акцепторы, а из ядер других — ядра As, т.е. мелкие доноры. Отношение количеств тех и других определяется сечениями соответствующих ядерных реакций и относительными концентрациями изотопов в облучаемом материале. При облучении природного германия получается p-Ge с коэффициентом компенсации $K = N_A/N_D = 0,4$. А сами концентрации N_A и N_D и концентрация носителей $n = N_A - N_D$ зависят только от времени облучения.

Для описания совокупности полученных кривых предполагается существование двух параллельных проводящих каналов с экспоненциальными температурными множителями в каждом:

$$\sigma = \sigma_b + \sigma_h = \sigma_{b0} \exp\left(-\frac{\varepsilon_b}{T}\right) + \sigma_{h0} \exp\left(-\frac{\varepsilon_h}{T}\right); \quad \varepsilon_b \gg \varepsilon_h. \quad (4.11)$$

Канал σ_b — это обычная зонная проводимость в полупроводниках с мелкими примесями, которая обусловлена тепловым забросом электронов с примесных уровней в зону проводимости и их делокализацией. Этот канал доминирует при более высоких температурах, когда сопротивление выходит на общую предельную прямую в левой части графика на рис. 4.1. Экспоненциальный рост сопротивления вдоль прямой

$$\rho = \rho_{b0} \exp \frac{\varepsilon_b}{T} \quad (4.12)$$

(ε_b это приблизительно энергия ионизации примеси) связан с вымерзанием носителей в зоне проводимости при низких температурах. Рост прекращается, когда основным становится другой механизм проводимости — прыжки электронов с одной примеси на другую без участия зоны делокализованных состояний. Температура кроссовера от одного механизма проводимости к другому, когда каналы σ_b и σ_h сравниваются, тем выше, чем больше концентрация N основных примесей (в примере на рис. 4.1 — это N_A).

В первую очередь обращает на себя внимание сильная зависимость величины σ_{h0} от концентрации примесей. Экстраполируя к $T = \infty$, получаем табл. 4.1.

Таблица 4.1

$10^{-15} N \text{ (см}^{-3}\text{)}$	0,15	1,5	3,5	35
$\sigma_{h0} (\Omega \cdot \text{см})^{-1}$	10^{-8}	10^{-6}	10^{-4}	10^{-1}

Это означает, что σ_{h0} имеет внутри себя еще один экспоненциальный множитель, так что экспериментальные данные описываются формулой

$$\sigma_h(T) = \sigma_{h0} \exp \left(-\frac{\varepsilon_h}{T} \right) = \sigma_{h0}^* \exp(-f(N)) \exp \left(-\frac{\varepsilon_h}{T} \right). \quad (4.13)$$

Сравнивая выражение (4.13) с зависимостями (4.8) и (4.10), получим

$$\sigma_h(T) = \sigma_{h0}^* \exp \left(-\frac{1,73}{N^{1/3} a_B} \right) \exp \left(-\frac{|\mu - E_D|}{T} \right). \quad (4.14)$$

Следующая из выражения (4.14) функциональная зависимость $\ln \rho \propto N^{-1/3}$ иллюстрируется двумя экспериментальными графиками — см. рис. 4.2. Из наклона экспериментальных прямых на этих графиках получается численный коэффициент в показателе первой экспоненты в выражении (4.14). Значения этого коэффициента из двух экспериментов рис. 4.2 вместе с другими аналогичными данными собраны в табл. 4.2.

Хотя в таблице собраны числа для разных материалов, они все близки к числу 1,73, фигурирующему в уравнении (4.10) и следующему из теории перколяции.

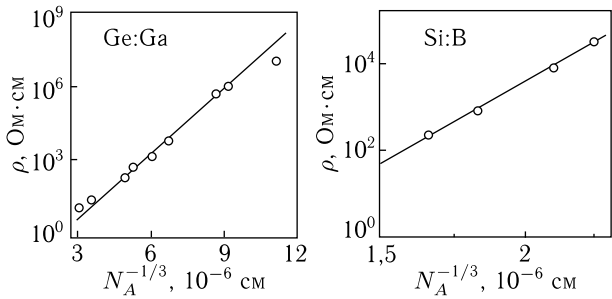


Рис. 4.2. Сопротивление в зависимости от среднего расстояния между акцепторами $N_A^{-1/3}$; Ge:Ga — из [3]; Si:B — из [4]

Таблица 4.2

n-GaAs	n-InP	p-Ge	p-Si
1,7	1,9	1,9	1,8
1,88		1,75	
1,9		2,0	

Самое важное в температурной зависимости на рис. 4.1 то, что эксперимент подтверждает функциональную зависимость $\ln(\rho/\rho_0) \propto 1/T$: экспериментальные зависимости в правой части графиков на рис. 4.1 — это прямые линии.

4.3. Прыжки с переменной длиной прыжка

Закон Мотта. Условие реализации проводимости путем прыжков на ближайшие центры состоит в том, что должно быть много таких пар близких соседей, в которых один центр свободен. Если понизить температуру, сделав

$$T \ll \mu - E_D, \tag{4.15}$$

то среди ближайших соседей, большинство которых имеют энергию E_D , окажется слишком мало свободных мест, и прыжки на ближайшие центры прекратятся. Тогда должны стать существенными прыжки между центрами, имеющими энергии в некоторой ε -окрестности ферми-уровня, где заведомо есть свободные места. Вопрос в том, сколь близко окажутся такие центры друг к другу.

Рассмотрим окрестность ферми-уровня $\mu \pm \varepsilon$ и будем считать плотность состояний в ней константой: $g = g_\mu$. Тогда число состояний в этой окрестности $N(\varepsilon) = g_\mu \varepsilon$, среднее расстояние между ними $\bar{r}_{ij} \approx [N(\varepsilon)]^{-1/3}$, а разность энергий порядка ε . Теперь основной статистический фактор F в уравнении (4.5) определяется вероятностью наличия фононов φ_{ij} , которые необходимы для выполнения закона

сохранения энергии. Поскольку энергия требуемого фонона $\Delta_{ij} \sim \varepsilon$, основной множитель в F порядка $\exp(-\varepsilon/T)$.

Оставим в случайной сетке Абрахамса–Миллера лишь те узлы, энергия которых попадает в интервал $\mu \pm \varepsilon$. Плотность узлов $N(\varepsilon)$ и среднее расстояние $\bar{r}_{ij}$ между ними в получившейся подсетке зависят от пока не определенной величины ε . Соседи в этой подсетке соединены сопротивлениями (4.6), в которых величина u_{ij} равна

$$u_{ij} = \frac{2}{a_B [N(\varepsilon)]^{1/3}} + \frac{\varepsilon}{T} = \frac{2}{g_\mu^{1/3} a_B \varepsilon^{1/3}} + \frac{\varepsilon}{T}. \quad (4.16)$$

Определим ε из условия, чтобы величина u_{ij} была минимальна:

$$\frac{d}{d\varepsilon} u_{ij}(\varepsilon) = 0, \quad \varepsilon = \varepsilon_{\min} = \left(\frac{T}{g_\mu^{1/3} a_B} \right)^{3/4} = (T^3 T_M)^{1/4}, \quad (4.17)$$

$$T_M \approx (g_\mu a_B^3)^{-1} > T.$$

Среднее изменение энергии при прыжках внутри так определенной подсетки $\varepsilon_{\min} \approx (T^3 T_M)^{1/4}$, а средняя длина прыжка

$$\bar{r} \approx \left(\frac{a_B}{g_\mu T} \right)^{1/4} \approx a_B \left(\frac{T_M}{T} \right)^{1/4}. \quad (4.18)$$

Вероятность прыжка на меньшее расстояние уменьшается за счет множителя $\exp(-\Delta_{ij}/T)$, а на большее — за счет множителя $\exp(-2r_{ij}/a_B)$. То, что средняя длина прыжка $\bar{r}$ зависит от T , оправдывает название этого типа проводимости.

Средняя длина прыжка $\bar{r} = \bar{r}_{ij}(\varepsilon_{\min})$ является радиусом взаимодействия (A.10) с точки зрения перколяционной теории. Хотя сама подсетка *меняется с температурой*, ее узлы, вероятности прыжков между которыми конечны,

$$u_{ij} \approx 2 \left(\frac{T_M}{T} \right)^{1/4},$$

образуют бесконечный кластер при любой температуре. Сопротивление подсетки равно

$$\rho = \rho_0 \exp \left(\frac{T_M}{T} \right)^{1/4}. \quad (4.19)$$

Величина ρ_0 зависит от температуры степенным образом, но обычно этой зависимостью пренебрегают.

Степень 1/4 в показателе экспоненты (4.19) получается только в трехмерном изоляторе. Для тонкой пленки

$$\bar{r}_{ij} \approx [N(\varepsilon)]^{-1/2}, \quad u_{ij} = \frac{2}{g_\mu^{1/2} a_B \varepsilon^{1/2}} + \frac{\varepsilon}{T}, \quad (4.20)$$

$$\varepsilon_{\min} = \left(\frac{T}{g_\mu^{1/2} a_B} \right)^{2/3} = (T^2 T_M)^{1/3}, \quad T_M \approx (g_\mu a_B^2)^{-1}, \quad (4.21)$$

и вместо (4.18) и (4.19) получается

$$\bar{r} \approx \left(\frac{a_B}{g_\mu T} \right)^{1/3} \approx a_B \left(\frac{T_M}{T} \right)^{1/3} \quad (4.22)$$

и

$$\rho = \rho_0 \exp \left(\frac{T_M}{T} \right)^{1/3}. \quad (4.23)$$

Используя обозначение d для размерности системы, формулы (4.19) и (4.23) можно объединить в одну:

$$\rho = \rho_0 \exp \left(\frac{T_M}{T} \right)^{1/(d+1)}, \quad T_M \approx (g_\mu a_B^d)^{-1}. \quad (4.24)$$

Закон Эфроса–Шкловского. При выводе закона Мотта плотность состояний $g(\varepsilon)$ вблизи ферми-уровня предполагалась постоянной. Но это не так при наличии кулоновской щели, когда

$$g(\varepsilon) \propto \left(\frac{\kappa}{e^2} \right)^d |\varepsilon|^{d-1}, \quad g(0) = 0 \quad (4.25)$$

(d — это размерность пространства, энергия ε отсчитывается от ферми-уровня — см. формулы (3.22) и (3.23) в гл. 3). Как и при выводе закона Мотта, введем симметричную относительно μ окрестность ферми-уровня $\mu \pm \varepsilon$. Число состояний в этой окрестности теперь зависит от размерности d и равно

$$N(\varepsilon) \propto \left(\frac{\kappa \varepsilon}{e^2} \right)^d.$$

Зато среднее расстояние $\bar{r}_{ij}$ между центрами, принадлежащими этой окрестности, от размерности не зависит:

$$\bar{r}_{ij} \approx [N(\varepsilon)]^{-1/d} \approx \frac{e^2}{\kappa \varepsilon}.$$

Теперь можно повторить все рассуждения предыдущего параграфа. Показатель экспоненты для величины сопротивления в подсетке Абрахамса–Миллера с разбросом энергий ε ,

$$u_{ij} = \frac{2}{a_B [N(\varepsilon)]^{1/d}} + \frac{\varepsilon}{T} = \frac{2e^2}{\kappa a_B \varepsilon} + \frac{\varepsilon}{T}, \quad (4.26)$$

имеет минимум при

$$\varepsilon_{\min} = (T T_{ES})^{1/2}, \quad T_{ES} \approx \frac{e^2}{\kappa a_B}, \quad (4.27)$$

и, независимо от размерности, средняя длина прыжка

$$\bar{r} \approx \left(\frac{e^2 a_B}{\kappa T} \right)^{1/2} \approx a_B \left(\frac{T_{ES}}{T} \right)^{1/2}, \quad (4.28)$$

а сопротивление меняется по закону

$$\rho = \rho_0 \exp \left(\frac{T_{ES}}{T} \right)^{1/2}. \quad (4.29)$$

Формулы (4.26)–(4.29) заменяют формулы (4.16)–(4.19) для трехмерной среды и формулы (4.20)–(4.23) для двумерной.

Заметьте. Имеющая размерность энергии величина T_{ES} не является шириной кулоновской щели $\delta\varepsilon$. Оценку для величины $\delta\varepsilon$ можно получить, приравняв величину (4.25) к плотности состояний на ферми-уровне $g(\mu)$, вычисленной без учета щели.

4.4. Экспериментальные наблюдения прыжковой проводимости

Таким образом теория предлагает несколько вариантов прыжковой проводимости. Задача эксперимента установить, где, когда и почему реализуется тот или иной вариант. Следует иметь в виду, что вариант прыжковой проводимости, реализующийся в конкретных условиях, зависит не только от соотношения параметров вещества, но и от температуры. На рис. 4.3 представлена естественная цепочка сменяющих

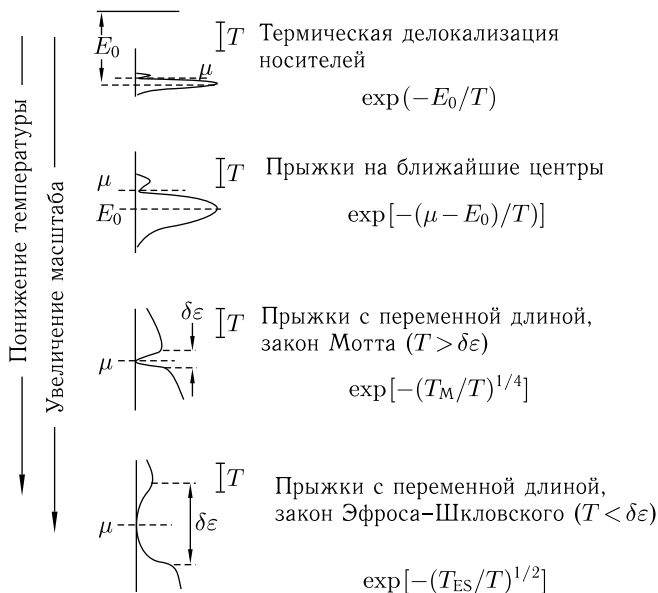


Рис. 4.3. Изменение механизма прыжковой проводимости по мере уменьшения температуры. При каждой температуре показаны особенности в плотности состояний в соответствующем масштабе.

друг друга с температурой механизмов проводимости. Эта цепочка построена в предположении, что ферми-уровень находится в примесной зоне локализованных состояний слабо легированного полупроводника. Каждый следующий более низкий уровень в цепочке соответствует более низкой температуре T , но на рис. 4.3 понижение T каждый

раз компенсируется увеличением масштаба, так что интервал энергий, равный T , остается неизменным.

Переход от зонной проводимости термически возбужденных носителей к прыжковой с прыжками на ближайшие примесные центры очень хорошо виден на примере p-Ge на рис. 4.1: это излом между прямой $\ln \rho \propto (1/T)$ в левой части графика, общей для всех концентраций N_A , и прямыми, подходящими к ней справа, из области более низких температур. Другие варианты смены доминирующего типа прыжковой проводимости тоже неоднократно наблюдались, но они не столь наглядны. Это связано со спецификой экспериментального определения степени температуры в показателе экспоненты. При построении экспериментального графика в координатах $(T^{-1/\nu}, \ln \rho)$, где $\nu = 1, 2, 3$ или 4, получается прямая, если степень ν подобрана правильно, и изогнутая линия в противном случае. Для того чтобы выяснить с достаточной точностью, укладывается ли получившаяся зависимость на прямую при той или иной степени T , отложенной вдоль оси абсцисс, нужен достаточно большой интервал изменения функции. Это иллюстрируют графики температурных зависимостей сопротивления n-InP. Отличить прыжки на ближайшие центры ($\nu = 1$) от прыжков с переменной длиной ($\nu \neq 1$) довольно просто. Проводимость при прыжках на ближайшие центры в координатах рис. 4.4, а должна изображаться прямой линией; пунктирная кривая на этом рисунке соответствует закону Мотта. Экспериментальные точки существенно ближе к кривой, чем к прямой. Различить варианты прыжков с переменной длиной труднее. В том же n-InP на рис. 4.4, б для того чтобы отличить $\nu = 1/2$ от $\nu = 1/4$, потребовался интервал изменения ρ в 5 порядков и скупуплезный учет предэкспоненциальных множителей.

Наиболее часто используемый и естественный диапазон температур в экспериментах по прыжковой проводимости это 4 — 0,04 К. То, какой именно механизм реализуется в этом диапазоне, зависит от конкретных параметров конкретного материала. Как видно из рис. 4.5, в GaAs в этом диапазоне температур $\ln \rho$ следует закону $T^{-1/2}$ без каких-то бы то ни было отклонений. Следовательно, в этом материале имеется кулоновская щель шириной порядка нескольких градусов. В отличие от этого на Si:As (рис. 4.6) сопротивление при некоторых концентрациях следует закону $T^{-1/4}$ от примерно 10 К по крайней мере до 0,5 К. Это означает, что в Si:As кулоновская щель имеет ширину, заведомо меньшую, чем 0,5 К, либо щели в спектре нет вовсе.

Все материалы, обсуждавшиеся до сих пор в этой главе, это классические полупроводники с мелкими донорами или акцепторами. Все эти материалы естественным образом становятся изоляторами при понижении температуры. Чтобы не возникло впечатления, что прыжковая проводимость — специфическое свойство этого весьма ограниченного класса веществ, приведем пример прыжковой проводимости в изоляторе совершенно другой природы — в пленке металлического вещества, столь тонкой, что проводимость в ней блокируется поверхностным рас-

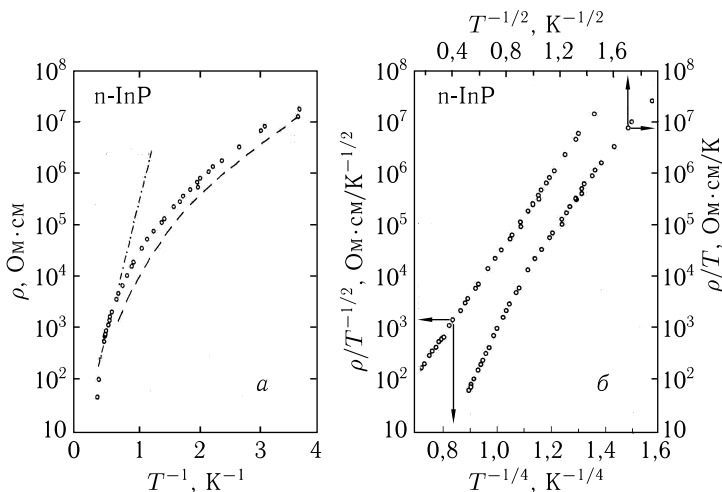


Рис. 4.4. Выяснение функциональной зависимости сопротивления n-InP от температуры путем построений в разных осях [5]. Штрихпунктирная прямая на рис. *a* показывает, как были бы расположены экспериментальные точки, если бы прыжки происходили между ближайшими соседями. Пунктирная линия соответствует температурным изменениям сопротивления при прыжковой проводимости по закону Мотта. На рис. *б* учтена температурная зависимость предэкспоненциального фактора

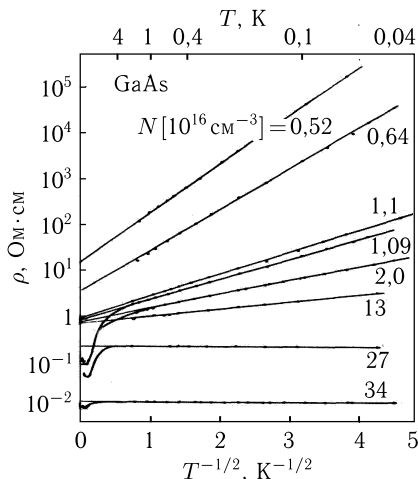


Рис. 4.5. График сопротивления образцов GaAs с различной концентрацией примесей в осях $(T^{-1/2}, \log \rho)$ [6]

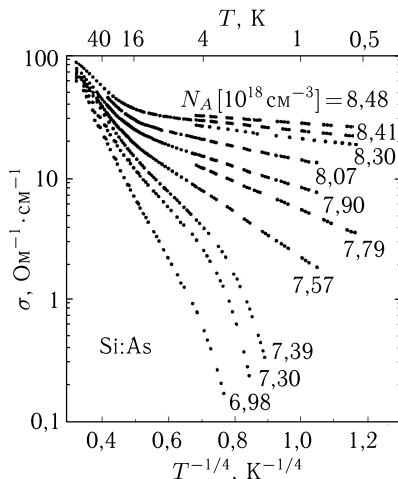


Рис. 4.6. График проводимости образцов Si:As с различной концентрацией примесей в осях $(T^{-1/4}, \log \rho)$ [7]

сеянием. Туннельные эксперименты на такой ультратонкой пленке Ве, выполненные при температурах $1 \div 0,5$ К, демонстрируют существование кулоновской щели (см. рис. 3.10 и рис. Б.5). Рис. 4.7 показывает, что эта пленка является изолятором и транспорт в ней осуществляется путем прыжковой проводимости. Как и должно быть при кулоновской щели, сопротивление меняется по закону $\ln \rho \propto T^{-1/2}$.

Таким образом экспериментально были обнаружены все описываемые теорией виды прыжковой проводимости. Тем не менее ряд фундаментальных вопросов пока остаются без экспериментальных ответов.

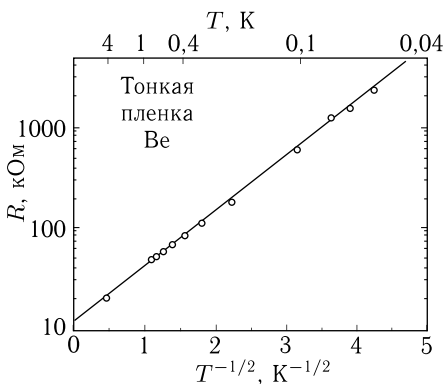


Рис. 4.7. Температурная зависимость сопротивления ультратонкой пленки Ве [8]

Во-первых остается неясным, всегда ли возникает кулоновская щель, или возможен изолятор с конечной плотностью состояний на ферми-уровне. То обстоятельство, что в экспериментах на n-InP (рис. 4.4) и на Si:As (рис. 4.6) проводимость изменяется по закону Мотта, строго говоря, еще не означает, что кулоновская щель в этих материалах вообще отсутствует. Установлена лишь верхняя граница $0,3 \div 0,5$ К возможной области существования кулоновской щели.

В принципе существование изолятора с локализованными состояниями на ферми-уровне и без кулоновской щели возможно. Как мы видели, для того чтобы образовалась щель, электроны должны иметь «пространство для маневра»: число вариантов распределения электронов по центрам должно быть достаточно велико. Этого может не быть из-за каких-то специфических особенностей структуры сети дефектов того или иного материала. Но экспериментальное изучение этого вопроса требует измерений при более низких температурах. Вместе с тем при расширении диапазона измерений в сторону низких температур из-за малости проводимости возникают затруднения с надежным ее измерением.

В трудность низкотемпературных измерений упирается изучение и второго фундаментального вопроса: с какой точностью выполняются в эксперименте уравнения (3.22) и (3.23) и подтверждается, что щель является мягкой. Согласно теоретическим оценкам, из-за кулоновского взаимодействия между несколькими частицами щель должна была бы стать жесткой. Отсюда вытекает естественная экспериментальная задача: сделать еще один шаг вниз на схеме на рис. 4.3 и при более низких температурах постараться определить структуру функции $g(\varepsilon)$ на расстояниях от ферми-уровня, существенно меньших, чем средняя ширина щели.

Следует иметь в виду еще одну возможность. Схема, изображенная на рис. 4.3, предполагает, что плотность состояний $g(\varepsilon)$ не зависит от температуры и что смена механизмов проводимости обусловлена величиной энергии, которая может быть передана в термостат. Однако возможно, что $g(\varepsilon)$ зависит от температуры и что мягкая щель превращается в жесткую при понижении температуры.

Все эти проблемы и трудности хорошо видны на примере изучения Si:B. График на рис. 4.8 с данными по сопротивлению Si:B в интервале температур от 4 К до 0,1 К иллюстрирует кроссовер от $1/4$ к $1/2$. Отклонение точек от прямой в нижней левой части основного графика с $T^{-1/2}$ на оси абсцисс показывает, что при температурах $T > 4$ К закон $1/2$ не работает. Но эта часть кривой спрямляется на вставке, построенная как функция от $T^{-1/4}$. Следовательно, выше 1 К работает закон Мотта, а ниже — закон Эфроса–Шкловского. Это согласуется с измерениями электронного спектра при помощи туннельного контакта, обнаружившими при 1,15 К кулоновскую щель в Si:B (см. рис. 3.9 в гл. 3 и рис. Б.5 в Приложении Б).

Однако экспериментальное наблюдение переходов с одного механизма на другой всегда страдает из-за того, что температурные интервалы при наличии кроссоверов недостаточно широки. Хотя самые низкотемпературные точки на основном графике на рис. 4.8 отклоняются от прямой в осях $(T^{-1/2}, \ln \rho)$, эти отклонения остаются в пределах ошибок. Однако, расширение интервала измерений в сторону низких температур (рис. 4.9) уточняет сделанные выводы. Закон Мотта при $T > 1$ К действительно выполняется. Однако при более низких температурах сопротивление Si:B ведет себя по-разному без магнитного поля и в сравнительно небольшом поле. В нулевом магнитном поле кроссовер происходит от $1/4$ не к $1/2$, а к 1, т. е. сопротивление при самых низких температурах следует закону Аррениуса $\rho \sim \exp(T_0/T)$ (рис. 4.9). Это означает, что кулоновская щель вблизи ферми-уровня из мягкой становится при более низких температурах жесткой: плотность состояний $g(\varepsilon)$ обращается в нуль не только в одной точке на ферми-уровне, но в некотором интервале T_0 энергий вокруг него; из графика следует, что $T_0 \approx 0,37$ К

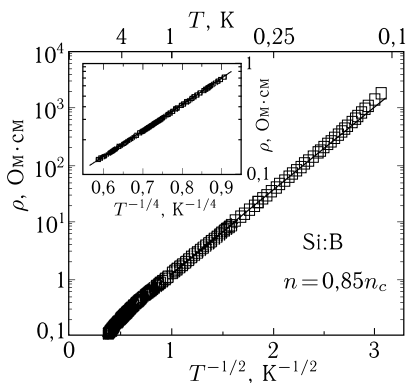


Рис. 4.8. Графики сопротивления образца Si:B в осях $(T^{-1/2}, \log \rho)$ и $(T^{-1/4}, \log \rho)$ [9]

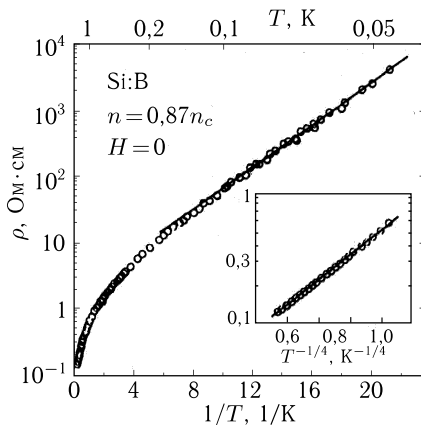


Рис. 4.9. Кроссовер в температурной зависимости сопротивления образца Si:B от закона Мотта $1/T$ к закону Аррениуса $1/T$. Минимальная температура ниже, чем на предыдущем рисунке [10]

(согласно туннельным измерениям, которые только что упоминались, ширина всей щели порядка 10 K).

Любопытно, что в магнитном поле щель остается мягкой: как видно из рис. 4.10, при $H > 2$ Тл точки ложатся на прямые вплоть до самых низких температур.

Неясно, стала ли щель жесткой из-за кулоновского или из-за какого-то иного взаимодействия. Зависимость от поля дает основания полагать, что щель становится жесткой из-за спиновых корреляций; мы не будем обсуждать здесь этот вопрос.

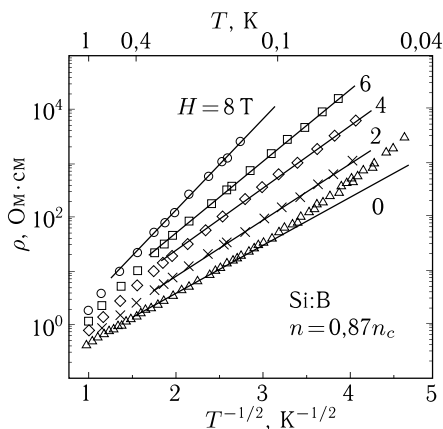


Рис. 4.10. Изменение функциональной зависимости сопротивления Si:B от температуры при наложении магнитного поля [10]

4.5. Контрольные вопросы и задачи

1. Во что превращается закон Мотта для трехмерного материала, если интеграл перекрытия убывает с расстоянием степенным образом, $J \propto (r_{ij}/a_B)^q$?

$$\rho \propto T^{-2q/3}.$$

2. Как по экспериментальным кривым прыжковой проводимости в р-Ge (рис. 4.1) оценить эффективный размер примесных центров a_B , по которым происходят прыжки?

3. Напишите формулы для прыжков по ближайшим соседям в случае двумерной системы, эквивалентные (4.9) и (4.14).

$$\pi r_c^2 N = 4,4; \quad \sigma \propto \exp(-1,18/N^{1/2} a_B).$$

4. Толщина b пленки п-типа со слабой компенсацией удовлетворяет неравенству $(N_D)^{-1/3} \ll b \ll (N_A)^{-1/3}$. Какова размерность системы с точки зрения описания прыжковой проводимости по ближайшим соседям; прыжковой проводимости с переменной длиной прыжка; формы кулоновской щели?

5. Согласно схеме на рис. 4.3, система при понижении температуры переходит из режима прыжков на ближайших соседей в режим закона Мотта, а затем в режим закона Шкловского–Эфроса. Может ли нарушиться этот порядок смены режимов прыжковой проводимости?

Список литературы

1. Шкловский Б. И., Эфрос А. Л. Электронные свойства легированных полупроводников. — Наука, 1979.
2. Hopping transport in solids (eds. M. Pollak, B. Shklovskii). — North-Holland, 1991.
3. Fritzsche H., Guevas M. // Phys. Rev. **119**, 1238 (1960).
4. Ray R., Fan H. // Phys. Rev. **121**, 768 (1961).
5. Mansfield R., Abboudy S., Foozoni F. // Philos. Mag. B **57**, 777 (1988).
6. Rentzsch R., Friedland K. J., Ionov A. N. et al. // Phys. Stat. Sol. b **137**, 691 (1986).
7. Shafarman W. N., Koon D. W., Castner T. G. // Phys. Rev. B **40**, 1216 (1989).
8. Butko V. Yu., DiTusa J. F., Adams P. W. // Phys. Rev. Lett. **84**, 1543 (2000).
9. Massey J. G., Lee M. // Phys. Rev. Lett. **75**, 4266 (1995).
10. Dai P., Zhang Y., Sarachik M. P. // Phys. Rev. Lett. **69**, 1804 (1992).

Глава 5

ПЕРЕХОДЫ МЕТАЛЛ–ИЗОЛЯТОР ¹⁾

Фундаментальное различие между изолятором и металлом в том, что в изоляторе электронные состояния на ферми-уровне локализованы, а в металле — делокализованы. Если последовательным изменением какого-то параметра перевести изолятор в металлическое состояние, то произойдет изменение симметрии волновых функций. Металл от изолятора отличается этой симметрией. Определив понятия металл и изолятор через волновые функции электронов, нужно сразу же отметить, что основное физическое свойство, принципиально различное в материалах этих двух типов, это величина проводимости, т. е. возможность проводить ток во сколь угодно слабом электрическом поле. Это признак типа «да–нет»: либо проводимость $\sigma = 0$, либо сколь угодно мала, но отлична от нуля. Однако при конечной температуре $T \neq 0$ изолятор тоже проводит ток путем прыжковой проводимости. Поэтому сформулированное определение изолятора относится только к температуре $T = 0$. Чтобы ответить на вопрос, является исследуемый материал металлом или изолятором, необходимо экстраполировать экспериментальную зависимость $\sigma(T)$ к $T = 0$, хотя эта процедура громоздкая, неудобная, а часто и неоднозначная.

Заметьте. Многие годы считалось, что металлы и изоляторы можно различать по знаку производной при низких температурах, где у металлов, мол, производная $\partial\sigma/\partial T \leq 0$, а у изоляторов $\partial\sigma/\partial T > 0$. Изучение квантовых поправок к проводимости показало, что $\partial\sigma/\partial T > 0$ может быть и у металлов, так что знак $\partial\sigma/\partial T$ не может быть критерием различия.

Поскольку наличие проводимости различает металл и изолятор только при $T = 0$, и под волновыми функциями, симметрия которых сравнивается, подразумеваются функции основного состояния, постольку само понятие перехода металл–изолятор имеет смысл только при $T = 0$. Известны два главных фактора, влияющие на волновые функции основного состояния, изменение которых может вызвать переход: степень беспорядка и электрон–электронное взаимодействие.

¹⁾ Обсуждение тех же вопросов, но в несколько ином стиле можно найти в книге [1] и в обзорах [2, 3].

Переход в системе невзаимодействующих электронов под влиянием изменения степени беспорядка называется переходом Андерсона. Переход, обусловленный межэлектронным взаимодействием, называется переходом Мотта. Обычно одновременно меняются оба фактора — тогда целесообразно говорить о переходе Андерсона–Мотта. Реальным управляющим параметром, который влияет на один или на оба ведущих фактора, может быть концентрация примесей, давление, магнитное поле и т.д. Обозначим значение управляющего параметра через x . На фазовой плоскости (x, T) переход изображается изолированной точкой на оси $T = 0$.

Очень важно. Здесь и дальше имеются в виду только чисто электронные фазовые переходы. Вместе с тем существуют также *структурные* переходы, сопровождающиеся изменением состояния атомной системы. Эти переходы мы рассматривать не будем. Заметим лишь, что уменьшение электронной энергии очень часто является основной причиной структурного фазового перехода. Таким является, например, переход Пайерлса. В модели, обсуждавшейся в гл. 3, переход в состояние изолятора происходит при конечной температуре, когда уменьшающийся с ростом температуры выигрыш электронной энергии от появления нового периода сравнивается с проигрышем упругой энергии, вызванным смещением ионов.

Структурные переходы очень разнообразны. Состояние электронной подсистемы по обе стороны такого перехода обычно оценивается на основании зонной теории, исходя из положения ферми-уровня относительно энергетических зон. Например, при температуре $T = 18^\circ\text{C}$ происходит переход белое олово — серое олово. Высокотемпературная фаза, белое олово, это хороший металл, в котором ферми-уровень пересекает несколько энергетических зон, образуя в них ферми-поверхности. В термодинамически стабильном при низких температурах сером олове ферми-уровень находится в запрещенной зоне; поэтому оно является изолятором. Мы будем называть такой изолятор зонным, чтобы отличать его от андерсоновского изолятора, в котором в непосредственной окрестности ферми-уровня есть электронные состояния, но лишь с локализованными волновыми функциями.

Взяв за основу какой-либо зонный изолятор, например, Ge, или Si, или то же серое олово, и вводя в него примеси, можно произвести в нем электронный переход в металлическое состояние. При этом для определения критической концентрации понадобятся низкотемпературные измерения.

Итак, под переходом металл-изолятор мы будем понимать только электронные переходы. Наша задача — обсудить, как эти переходы происходят при изменении степени беспорядка и электронной плотности.

5.1. Переход Андерсона

Рассмотрим вслед за Андерсоном трехмерную периодическую решетку ям различной глубины с концентрацией $N = a^{-3}$ (a — период решетки; см. рис. 5.1). Пусть уровни в ямах разбросаны в интервале значений энергии W и плотность уровней в этом интервале постоянна. Выбрав точку отсчета энергии $\varepsilon = 0$ в середине интервала W , получим:

$$\frac{\partial N}{\partial \varepsilon} = \begin{cases} N/W & \text{при } |\varepsilon| \leq W/2, \\ 0 & \text{при } |\varepsilon| > W/2. \end{cases} \quad (5.1)$$

Заметьте. Здесь уровень $\varepsilon = 0$ — это не ферми-уровень. Количество электронов в системе зависит от внешних обстоятельств.

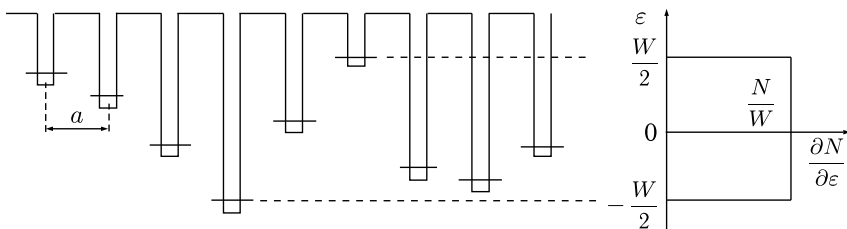


Рис. 5.1. Модель Андерсона: периодически расположенные ямы разной глубины

Благодаря хвостам волновых функций $\exp(-r/a_B)$ существует перекрытие волновых функций электронов, локализованных на соседних ямах (ср. формулы (4.1) и (4.2) в начале гл. 4). Если расстояние между соседними ямами $r_{12} \gg a_B$, то интеграл перекрытия

$$J = \int \psi_1^* \hat{H} \psi_2 d^3r \propto \exp\left(-\frac{r_{12}}{a_B}\right) \quad (5.2)$$

мал и его малость определяется множителем $\exp(-r_{12}/a_B)$.

В принципе возможны два предельных случая. Каждый электрон может сидеть в своей яме; так будет, например, если ямы хоть и разные, но все очень глубокие. С другой стороны, все электроны могут быть делокализованы, так что каждому электрону доступна любая яма. Например, если все ямы одинаковые или почти одинаковые, то волновые функции электронов это блоховские волны.

Утверждение Андерсона, сформулированное им для трехмерного множества периодически расположенных ям, звучит так: симметрия волновых функций основного состояния зависит от относительной величины беспорядка. Параметром является отношение энергий J/W .

Для появления делокализованных состояний, т. е. для реализации металлической проводимости, требуется выполнение условия

$$\frac{J}{W} \geq \left(\frac{J}{W} \right)_{\text{crit}}. \quad (5.3)$$

При критическом значении отношения J/W делокализованные состояния появляются в центре зоны при $\varepsilon = 0$, дальнейшее увеличение отношения J/W приводит к постепенному расширению слоя делокализованных состояний.

Смысл и роль отношения J/W можно пояснить при помощи простейших задач квантовой механики. Конечное перекрытие двух ям различной глубины с уровнями E_{10} и E_{20} и волновыми функциями φ_1 и φ_2 приводит в первом порядке теории возмущений к малым поправкам к волновым функциям

$$\psi_1 = c_1 \varphi_1 + c_2 \varphi_2, \quad \psi_2 = -c_2 \varphi_1 + c_1 \varphi_2, \quad c_2 = \frac{J}{E_{10} - E_{20}}. \quad (5.4)$$

Поскольку $c_2 \ll c_1 \approx 1$, каждый электрон остается в основном в своей яме.

Если ямы одинаковые, «резонансные», $E_{10} = E_{20} = E_0$, то ответ принципиально другой: уровень E_0 расщепится на $E_{1,2} \approx E_0 \pm J$ (рис. 5.2, б), а волновые функции в обоих состояниях размажутся равномерно по обеим ямам:

$$\psi_{1,2} = \frac{1}{\sqrt{2}} (\varphi_1 \pm \varphi_2). \quad (5.5)$$

Различаются не только волновые функции (5.4) и (5.5), но и порядки величин сдвигов уровней (рис. 5.2). Каждая из двух ям различной

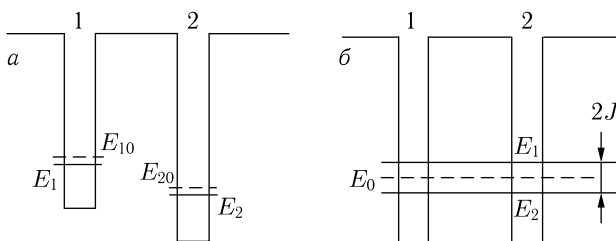


Рис. 5.2. Сдвиг уровней в двух прямоугольных ямах разной глубины (а) и расщепление уровней в двух одинаковых ямах (б) при учете перекрытия в первом порядке теории возмущений

глубины является возмущением для волновой функции электрона в другой яме. Поскольку невозмущенная волновая функция φ_1 в окрестности ямы 2 имеет порядок $\exp(-r_{12}/a_B)$, сдвиг уровня E_{10} порядка

$$\Delta_1 E \equiv E_1 - E_{10} \approx \int \varphi_1^* \hat{H}_2 \varphi_1 d^3 r \propto \exp\left(-\frac{2r_{12}}{a_B}\right), \quad (5.6)$$

т. е. малый множитель $\exp(-r_{12}/a_B)$ из интеграла перекрытия входит в $\Delta_1 E$ в квадрате (рис. 5.2, а). Для резонансных ям (рис. 5.2, б)

$$\Delta E \propto \exp\left(-\frac{r_{12}}{a_B}\right), \quad \text{т. е.} \quad \Delta E \sim J. \quad (5.7)$$

Ямы ведут себя как резонансные пока разность невозмущенных энергий у них $|E_{10} - E_{20}| < J$. Следовательно J/W — это доля резонансных узлов. Тогда качественно критическое значение $(J/W)_{\text{crit}}$ можно трактовать как перколяционный порог, при достижении которого в спектре появляются состояния с делокализованными волновыми функциями.

Если значение параметра J/W меньше критического и волновые функции на ферми-уровне локализованы, то вещество является изолятором. Такой изолятор называется андерсоновским. Он устроен принципиально иначе, нежели зонный. В зонном изоляторе ферми-уровень расположен в запрещенной зоне, где плотность состояний равна нулю, а проводимость обеспечивается за счет электронов, термически заброшенных в зону проводимости, или за счет дырок, заброшенных в валентную зону. В андерсоновском изоляторе плотность состояний на ферми-уровне конечна. Уровень энергии, начиная с которого состояния становятся делокализованными, находится на некотором расстоянии от ферми-уровня. Он называется край подвижности и играет роль дна зоны проводимости. Электроны или дырки, термически заброшенные за край подвижности, участвуют в проводимости путем диффузии, дырки или электроны ниже края подвижности участвуют в прыжковой проводимости. Переход в металлическое состояние происходит при совмещении ферми-уровня и края подвижности. Этого можно достичь, либо меняя параметр J/W , либо двигая ферми-уровень.

5.2. Модель структурного беспорядка ¹⁾

Весьма поучительна другая одноэлектронная модель беспорядка, отличающаяся от модели Андерсона. В этой модели случайный потенциал состоит из одинаковых, но хаотично расположенных ям, в каждой из которых имеется уровень E_0 :

$$V(\mathbf{r}) = \sum_{\mathbf{R}_i} v(\mathbf{r} - \mathbf{R}_i). \quad (5.8)$$

Беспорядок в этой модели определяется хаотичностью множества векторов $\mathbf{R}_i$.

¹⁾ В книге [4] эта модель фигурирует под названием модель Лифшица; ее математические аспекты подробно обсуждаются в обзоре самого Лифшица [5], который применил эту модель для исследования энергетического спектра металла с примесями.

Обратите внимание на качественно разный характер беспорядка в модели Андерсона (5.1) и в модели структурного беспорядка (5.8). Можно сказать, что модель Андерсона происходит из физики полупроводников, где энергии примесных центров (см. ниже рис. 5.4) всегда различны, а модель структурного беспорядка родом из физики металлов, где благодаря сильному экранированию энергии всех центров можно считать одинаковыми. Не случайно потенциал (5.8) используется в дифракционной теории электронного транспорта в жидких металлах (см. гл. 1). Там предполагается, что все электроны делокализованы, а потенциал (5.8) является лишь источником рассеяния. Это соответствует большому отношению длины затухания a_B к среднему расстоянию между ямами $n^{-1/3}$, т. е. большому значению параметра $a_B n^{1/3}$.

Предполагается по-прежнему, что за пределами ямы волновая функция затухает по закону

$$\psi \propto \frac{e^{-r/a_B}}{r}, \quad a_B = \frac{\kappa \hbar^2}{m^* e^2}. \quad (5.9)$$

Пусть сначала концентрация мала, так что длина затухания гораздо меньше среднего расстояния между центрами: $a_B n^{1/3} \ll 1$. Тем не менее за счет слабого перекрытия волновых уровней соседей изначально δ -образная плотность состояний расплывается в полосу. Разобьем все ямы на пары ближайших соседей. Если в такой паре расстояние между ямами r_{12} , то, поскольку ямы резонансные, из-за перекрытия хвостов (5.9) волновых функций в ямах возникают два уровня с энергиями

$$E = E_0 \pm \varepsilon_{1,2}, \quad \varepsilon_{1,2} = J_0 \frac{\exp(-r_{12}/a_B)}{r_{12} n^{1/3}} \quad (5.10)$$

и с обобществленными волновыми функциями (5.5). В выражении для расщепления уровней (5.10) расстояние r_{12} нормировано на среднее расстояние между ямами $n^{-1/3}$ для того, чтобы J_0 имело размерность энергии. Величина J_0 , как и величина a_B , зависит от конкретных характеристик ямы, диэлектрической проницаемости материала, эффективной массы электрона и т. д.

Как видно из рис. 5.3, где резонансные пары показаны пунктирными эллипсами, разбить на резонансные пары можно далеко не все центры. Например, яма 2, будучи ближайшим соседом ямы 3, может иметь своим ближайшим соседом яму 1, так что $r_{12} < r_{23}$. В этой тройке наибольшими будут резонансные сдвиги энергий ε_1 и ε_2 , а сдвиг ε_3 будет нерезонансным и существенно меньшим, так как $\varepsilon_3 \propto \exp(-2r_{23}/a_B)$. На рис. 5.3 показаны две такие конфигурации, одна тройка и одна четверка, в которой $r_{12} < r_{23} < r_{34}$. Но и в тройках, и в более сложных конфигурациях из четырех и более ям всегда имеется по крайней мере одна резонансная пара с наименьшим расстоянием между ямами и наибольшим сдвигом уровней. Характерная ширина Δ получающейся

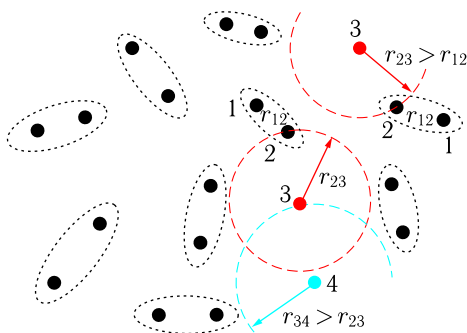


Рис. 5.3. Случайное расположение примесных ям. Пары ближайших соседей помечены пунктирными эллипсами. Центры с индексом 3, примыкая к парам, ямы которых помечены индексами 1 и 2, образуют с ними тройки

в результате функции плотности состояний определяется резонансными парами ям, а среднее расстояние между ямами равно $n^{-1/3}$, так что

$$\Delta \approx J_0 \exp \left(-\frac{n^{-1/3}}{a_B} \right). \quad (5.11)$$

Хвост плотности состояний при $|\varepsilon| \gg \Delta$ получается за счет пар anomalно близко расположенных ям с $r_{12} \ll n^{-1/3}$, состояния при малых $|\varepsilon| \ll \Delta$ — за счет нерезонансных и уединенных ям.

Поучительно сравнить описанное превращение в полосу изначально δ -образной плотности состояний с формированием примесной зоны в системе «доноры + акцепторы» (см. гл. 3). Там расплывание происходило за счет случайных электрических полей заряженных центров при частичной компенсации примесей. Такое расширение уровня в полосу можно назвать классическим. А уже в классически уширенной примесной зоне переход металл–изолятор происходит за счет перекрытия уровней при дальнейшем увеличении концентрации примесных центров. В отличие случая, от этого уширение (5.11) имеет квантовую природу. Параметр $a_B n^{1/3}$ контролирует и превращение уровня в полосу, и переход.

Как видно из формулы (5.11), в модели структурного беспорядка нет независимого параметра, эквивалентного ширине полосы W в модели Андерсона (5.1)–(5.3). Возможно именно поэтому для описания переходов металл–изолятор чаще используется именно модель (5.1)–(5.3). На рис. 5.1 случайные классические поля, всегда имеющиеся в реальных системах, в явном виде не фигурируют; они как бы включены в дисперсию глубин ям. Но их можно изобразить и в явном виде, заменив рис. 5.1 на практически эквивалентный ему рис. 5.4. После такой замены он лучше отображает реальную ситуацию, например в частично компенсированном полупроводнике, где все примесные центры химически одинаковы.

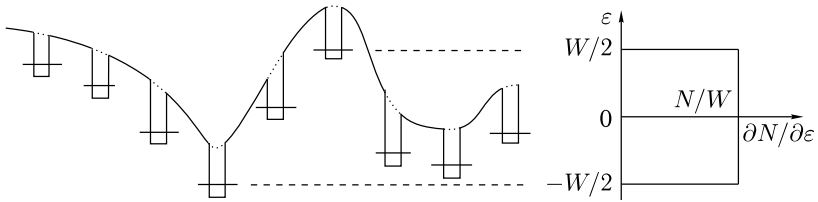


Рис. 5.4. Альтернативное представление модели Андерсона: периодически расположенные ямы одинаковой глубины на фоне случайного потенциала (ср. рис. 5.1)

Переход в модели структурного беспорядка теоретически исследован гораздо менее детально, чем в модели Андерсона. Однако при помощи численного моделирования было продемонстрировано существование перехода по параметру $a_B n^{1/3}$. Но помимо естественных управляющих параметров: формы и глубины ямы $v(\mathbf{r} - \mathbf{R}_i)$ и концентрации ям N , в модели структурного беспорядка есть и скрытый механизм влияния на переход. Такое влияние может реализовываться через величину корреляций на множестве векторов $\mathbf{R}_i$ — увеличивая корреляции, можно превратить это множество из хаотического в регулярное. Подобная возможность, по-видимому, реализуется в некоторых металлических расплавах и в квазикристаллах (см. гл. 7).

5.3. Переход Мотта

Рассмотрим систему доноров с водородоподобными волновыми функциями электронов (5.9). Сравним боровский радиус a_B с другой величиной, имеющей размерность длины, — с радиусом экранирования r_e электрического поля, который используется при описании системы свободных носителей. При статистике Ферми в r_e входит концентрация электронов n :

$$r_e = \left(\frac{4m\epsilon^2 n^{1/3}}{\kappa \hbar^2} \right)^{-1/2} = \frac{1}{2} (a_B n^{-1/3})^{1/2}. \quad (5.12)$$

Если все электроны локализованы, то система описывается длиной a_B , если делокализованы, то длиной r_e . Кроме того, в обоих случаях фигурирует еще и третья длина — среднее расстояние между донорами или между носителями $n^{-1/3}$. Предположим, что мы постепенно увеличиваем концентрацию доноров n . Пока $r_e > a_B$, экранирование несущественно, каждый электрон находится около своего донора, вещество является изолятором. Но когда соотношение между r_e и a_B изменится на $r_e < a_B$, то состояние изолятора станет неустойчивым. Если все n электронов уйдут от своих доноров, то они к ним не смогут вернуться, потому что не найдут их из-за сильного экранирования. Ионизованные

доноры создадут положительный фон, компенсирующий отрицательный заряд делокализованных электронов. Поэтому равенство

$$r_e = a_B, \quad \text{т. е.} \quad \frac{1}{2}(a_B n_c^{-1/3})^{1/2} = a_B, \quad a_B n_c^{1/3} = 0,25 \quad (5.13)$$

есть условие перехода металл–изолятор, который происходит, когда концентрация n достигает критического значения n_c . Такой переход называется переходом Мотта.

Заметьте. В этом нехитром рассуждении беспорядок никак не фигурирует: движущей пружиной перехода Мотта является межэлектронное взаимодействие. Можно мысленно представить себе, что доноры упорядочены в сверхрешетку, а концентрацию мы меняем, меняя период сверхрешетки.

При написании соотношений (5.12) и (5.13) использовано выражение для боровского радиуса (5.9), справедливое лишь для водородоподобного примесного центра. В качестве исходной точки для построений теории перехода Мотта обычно используется несколько иное, более общее рассуждение.

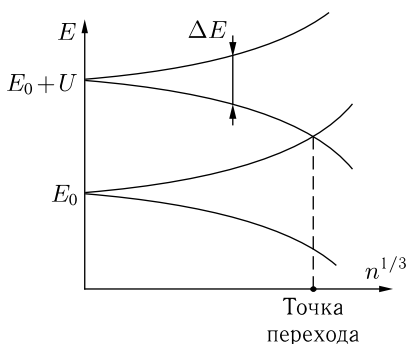


Рис. 5.5. Переход Мотта в схеме Хаббарда

Пусть энергия электрона на примесном центре E_0 . Принцип Паули допускает пребывание на этом уровне двух электронов с разными спинами. Но если один электрон на центре уже есть, то посадить туда второй можно только преодолев электростатическое отталкивание. Следовательно уровень для второго электрона будет выше на величину порядка $U \approx e^2/\kappa a_B$. Величина U называется энергией Хаббарда. Если концентрация

примесных центров $n \neq 0$, то, независимо от того, расположены центры хаотически или регулярно, из-за перекрытия волновых функций оба уровня размоются в минизоны с nV уровнями в каждой (V — объем). Поскольку концентрация электронов тоже равна n , то, если минизоны не перекрываются, все nV уровней в нижней минизоне заполнены, а в верхней пустые, и вещество является изолятором со щелью на ферми-уровне (см. рис. 5.5).

Заметьте. На любом центре в нижней зоне Хаббарда может находиться только один электрон, но с любым направлением спина. Поэтому получившийся изолятор будет парамагнетиком. Ферро- или антиферромагнитное упорядочение в этом изоляторе может возникнуть только при наличии дополнительного взаимодействия между центрами.

Ширина минизон ΔE определяется интегралами перекрытия (5.2):

$$\Delta E \approx 2J \approx 2 \int \psi_i^* \hat{H} \psi_j d^3r.$$

Оценка интегралов J базируется на выражении (5.9) для асимптотики волновых функций. Интеграл $\int \psi_j^* \hat{H} \psi_j d^3r$ определяет невозмущенный уровень энергии и, следовательно, равен E_0 . Замена ψ_j^* на ψ_i^* означает, что везде, где подынтегральное выражение существенно отлично от нуля, оно уменьшено в $\exp(-r_{ij}/a_B)$ раз. Поэтому

$$\Delta E \approx 2J \approx 2 \int \psi_i^* \hat{H} \psi_j d^3r \approx 2E_0 \exp\left(-\frac{r_{ij}}{a_B}\right) \approx 2E_0 \exp\left(-\frac{1}{a_B n^{1/3}}\right). \quad (5.14)$$

Увеличение концентрации приводит к расширению зон и к их перекрытию. Зоны сомкнутся при критической концентрации n_c , когда

$$U \approx \Delta E \approx 2E_0 \exp\left(-\frac{1}{a_B n^{1/3}}\right). \quad (5.15)$$

Поскольку по порядку величины $U \approx E_0$, численное значение параметра $a_B n^{1/3}$ на переходе вряд ли сильно отличается от 0,25, фигурировавшего в соотношении (5.13).

Таким образом, мы познакомились с двумя рассуждениями, приводящими к переходу металл-изолятор. Сравним их между собой и с двумя типами квантовых поправок к проводимости грязных металлов при низких температурах (гл. 2). Андерсоновский переход происходит вследствие беспорядка, причем для его описания достаточно одноэлектронного приближения. То же самое можно сказать и о слабой локализации. Напротив, в переходе Мотта движущей силой является кулоновское взаимодействие, регулируемое и изменяемое через экранирование, а степень беспорядка существенна в гораздо меньшей степени. Для межэлектронной интерференции беспорядок нужен лишь для того, чтобы движение электронов было диффузионным, а сама квантовая поправка к проводимости есть результат электрон-электронного взаимодействия.

Для экспериментатора представляется очень важным определить-ся, с каким типом перехода, Мотта или Андерсона, он имеет дело. Но практически почти всегда выбор оказывается не вполне убедительным. Изменение электронной концентрации n возможно лишь при нарушении стехиометрии или внесении примесей. Поэтому параллельно с изменением n меняется и беспорядок. С другой стороны, рост беспорядка влияет на экранирование. Количественные характеристики переходов Мотта и Андерсона тоже, как это ни странно, сходны. Уравнение перехода Андерсона (5.3) можно записать в виде

$$a_B n^{1/3} = -\left(\ln \frac{c_a W}{E_0}\right)^{-1}, \quad c_a = \left(\frac{J}{W}\right)_{\text{crit}}. \quad (5.16)$$

Уравнение перехода Мотта (5.15) можно записать в аналогичном виде

$$a_B n^{1/3} = - \left(\ln \frac{c_m U}{E_0} \right)^{-1}, \quad (5.17)$$

введя численную константу c_m .

Вспомните. Формулы для двух типов квантовых поправок тоже очень похожи друг на друга.

Величины W , U и E_0 , входящие в буквенные оценки констант, все порядка атомной энергии $e^2/\kappa a_B$, а численные множители c_a и c_m порядка единицы. К тому же все эти величины входят под знаком логарифма. В результате оценки для произведения $a_B n_c^{1/3}$ для обоих переходов практически одинаковы (на рис. 5.6 эта «похожесть» изображена в виде диаграммы). Поэтому не должна вызывать удивление сводка экспериментальных данных по различным переходам, представленная на графике рис. 5.7. Пропорциональность $a_B \propto n_c^{-1/3}$ подтверждается на интервале изменений n_c в восемь порядков.

ПЕРЕХОД МОТТА

Ширина ΔE

пропорциональна

$$J = E_0 \exp(-1/a_B n^{1/3})$$

$$\Delta E = U$$

$$a_B n^{1/3} = -(\ln c_m U / E_0)^{-1}$$

$$a_B n_c^{1/3} = 0,25$$

Соотношение Мотта

$$a_B n^{1/3} = -(\ln c_a W / E_0)^{-1}$$

Интеграл
перекрытия

$$J = E_0 \exp(-1/a_B n^{1/3})$$

входит непосредственно

$$W/J = \text{const}$$

ПЕРЕХОД АНДЕРСОНА

Рис. 5.6. Сходство и различия в выражениях, описывающих переходы Мотта и Андерсона

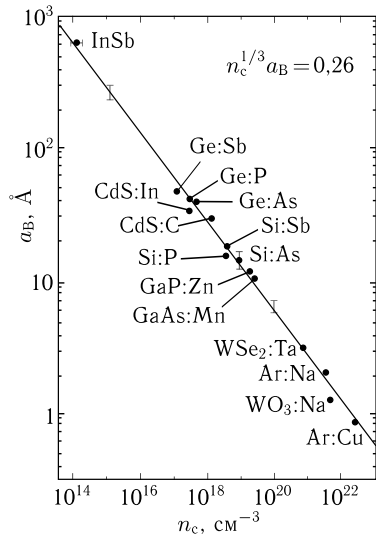


Рис. 5.7. Корреляция между критической концентрацией носителей и эффективным боровским радиусом при переходах металл-изолятор в 15 различных материалах. Для всех материалов величины a_B и n_c определялись независимо из разных экспериментов [6]

Заметьте. Поскольку критерий $a_B n^{1/3} = \text{const}$ справедлив для обоих типов переходов, то концентрация n может быть управляющим параметром не только для перехода Мотта (что естественно), но и для перехода Андерсона, в котором межэлектронное взаимодействие не фигурирует. При этом величина n выступает в роли концентрации не электронов, а центров их локализации, расстояние между которыми определяет величину интеграла перекрытия.

Заметьте также, что в модели структурного беспорядка в качестве условия того, что все носители локализованы, предполагалась малость того же самого параметра $a_B n^{1/3}$.

Для иллюстрации соотношения между переходами Мотта и Андерсона обратимся к рис. 5.8, на котором представлена плоскость (беспорядок–концентрация). В качестве меры беспорядка выбрана величина W , по вертикальной оси отложена концентрация. Будем считать значение боровского радиуса a_B постоянным. Справа от линии переходов Андерсона (5.16) расположены состояния изолятора, а слева — металла. Линия переходов Мотта (5.17) в этих осях описывается горизонтальной прямой, расположенной тем выше, чем меньше a_B . Линии пересекаются в точке $W_0 = (c_m/c_a)U$. Пока беспорядок большой, $W > W_0$, переходы металл–изолятор контролируются им и происходят вдоль линии (5.16). Малый беспорядок $W < W_0$ несущественен, потому что за счет межэлектронного взаимодействия локализация происходит при концентрациях n , больших, чем те, что следуют из (5.16).

Беспорядок, создаваемый путем смещения атомов, имеет некоторый верхний предел. Этот предел достигается, когда корреляции между положениями атомов вообще отсутствуют. На диаграмме на рис. 5.8 этот предел обозначен как W_{max} на правом краю полосы возможных состояний. Беспорядку W_{max} на кривой переходов соответствует концентрация n_{max} . Возникает вопрос, каковы концентрации электронов n в настоящих металлах и сплавах. Если $n > n_{\text{max}}$, то никаким беспорядком не удастся реализовать переход Андерсона.

Заметьте. Величина W может считаться количественной мерой беспорядка лишь весьма условно. Поэтому диаграмма на рис. 5.8

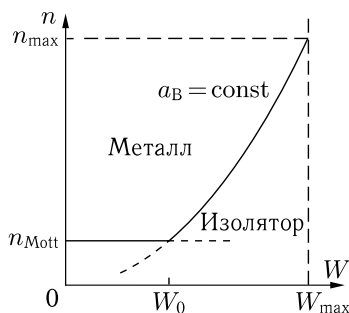


Рис. 5.8. Диаграмма (беспорядок–концентрация), на которой в качестве меры беспорядка для иллюстрации выбрана величина W , а величина a_B предполагается константой. Наклонная кривая — линия переходов Андерсона, горизонтальная прямая — линия переходов Мотта

носит иллюстративный характер. Тем не менее вопрос существует и ответить на него могут только эксперименты. Такие эксперименты уже обсуждались в гл. 1. Мы вернемся к этому вопросу в гл. 7.

5.4. Минимальная металлическая проводимость?

Предложенные критерии перехода в форме величины параметра $a_B n_c^{1/3}$ постулируют сам факт существования перехода металл–изолятор, но ничего не говорят о его характере. Первые серьезные дискуссии на эту тему базировались на анализе природы металлической проводимости. Формула для удельной проводимости

$$\sigma = \frac{e^2}{\hbar} \frac{nl}{k_F}, \quad n = \frac{1}{3\pi^2} k_F^3, \quad (5.18)$$

исходит из представления об электронной системе как о газе заряженных частиц с длиной свободного пробега l . Как и в обычном газе, на каком-то этапе возникают ограничения, обусловленные конечностью размеров частиц. Эффективным размером электронов является де-бройлевская длина волны $1/k_F$. Это естественное ограничение снизу для величины l (правило Иоффе–Регеля). При длине пробега $l \sim 1/k_F$ формула для σ принимает вид

$$\sigma_{\text{Mott}} \approx (3\pi^2)^{-2/3} (e^2/\hbar) n^{1/3} \approx 0,1(4 \cdot 10^3 \text{ Ом})^{-1} n^{1/3}. \quad (5.19)$$

Это значение ограничивает снизу проводимость, возможную в рамках газовой модели.

Исходя из этого ограничения, Мотт предположил, что переход металл–изолятор не является непрерывным фазовым переходом. Для перехода Андерсона рассуждения Мотта звучат так: по мере увеличения беспорядка в системе с фиксированной концентрацией n проводимость падает до минимума (5.19), а затем должна скачком обратиться в нуль. А уж при переходе Мотта сам сценарий предполагает лавинообразное увеличение числа делокализованных носителей: отдельный электрон, делокализовавшись, через свой вклад в экранирование способствует делокализации остальных.

Более того, базовое утверждение, лежащее в основе концепции о переходе Мотта, гласит, что концентрация n не может быть сколь угодно малой. Подстановка n_c из критериев (5.13) или (5.15) в выражение (5.19) выражает σ_{Mott} через a_B . Но a_B не может быть макроскопически большим — ведь это характерный размер волновой функции отдельного электрона. Значение $a_B \approx 600 \text{ \AA}$ в InSb (см. рис. 5.7), вероятно, близко к максимально возможному. Отсюда должно следовать существование абсолютной минимальной металлической проводимости.

Естественно, что концепция минимальной металлической проводимости и характер перехода металл–изолятор были подвергнуты тщательной проверке в многочисленных экспериментах.

Пример одного такого эксперимента приведен на рис. 5.9. Зависимость проводимости образцов Si:P от концентрации n (на вставке) из-за неточности определения концентрации в каждом отдельном образце может лишь определить критическую концентрацию n_c , но не закон изменения проводимости вблизи нее. Однако прикладывая давление к образцу с концентрацией, чуть меньшей n_c , оказалось возможно пройти интервал значений σ от нуля до предполагаемого σ_{Mott} и убедиться в отсутствии скачка проводимости.

Заметьте. Каждая точка на графике $\sigma(S)$ получена экстраполяцией к $T = 0$ температурной зависимости $\sigma(T)$, измеренной при данном давлении. Подобный график при любой конечной температуре был бы недостаточно убедителен, потому что в изоляторе при конечной температуре существует конечная проводимость.

Из описанного эксперимента следует, что в окрестности перехода металл–изолятор в Si:P *проводимость может принимать сколь угодно малые значения, так что переход является непрерывным.* По-видимому, так ведут себя все системы с критической концентрацией носителей $n_c \lesssim 10^{19} \text{ см}^{-3}$, т.е. системы на базе легированных

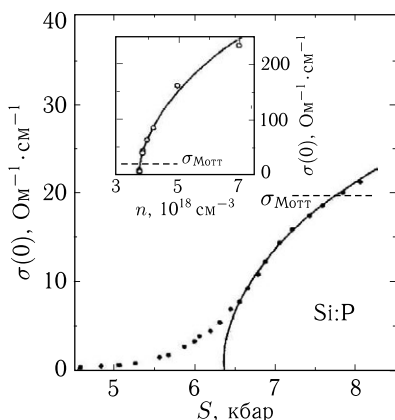


Рис. 5.9. Тонкая настройка давлением перехода металл–изолятор в Si:P. Пунктиром на основном графике и на вставке отмечена минимальная проводимость σ_{Mott} , рассчитанная по измеренной электронной концентрации [7]

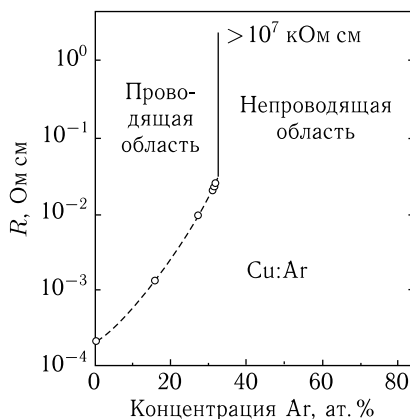


Рис. 5.10. Сопротивление при температуре около 4 К смеси медь–аргон, напыленных на холодную подложку, в зависимости от состава смеси [8]

полупроводников (см. рис. 5.7). Кривая $\sigma(S)$ ставит вопрос о модели проводимости в интервале

$$0 < \sigma < \sigma_{\text{Mott}}, \quad (5.20)$$

где не работают ни газовая модель, ни модель прыжковой проводимости. Мы постараемся дать на него качественный ответ в следующей главе, посвященной скейлинговой теории непрерывных квантовых переходов металл–изолятор.

На рис. 5.10 приведен результат «контр-эксперимента» [8] на неупорядоченной смеси атомов Ag и Cu, полученной одновременным напылением этих атомов на холодную подложку. Меняя время напыления и относительную интенсивность потоков атомов Cu и Ag, можно было регулировать толщину пленки Ag:Cu и относительную концентрацию Cu, а многочисленные калибровочные и измерительные процедуры позволяли определять эти величины с достаточной точностью. У многочисленных пленок с концентрацией Cu, близкой к критической, получалось удельное сопротивление либо в районе $3 \cdot 10^{-2}$ Ом см, либо по крайней мере на семь порядков больше, но никогда в промежутке. Это, а также и аналогичный результат на смеси Ag:Na, отмеченный на рис. 5.7, заставляют предполагать, что при концентрациях $n_c \gtrsim 10^{21} \text{ см}^{-3}$ проводимость на переходе металл–изолятор *обращается в нуль скачком*. Однако, это требует проверки на других системах с концентрацией электронов в интервале $10^{20} \text{ см}^{-3} \lesssim n \lesssim 10^{22} \text{ см}^{-3}$.

Отметим в заключение, что на качественной диаграмме на рис. 5.8 увеличение параметра $a_B = \text{const}$ приводит к смещению прямой $n = n_{\text{Mott}}$ вверх, что может привести к качественному изменению диаграммы. К тому же напомним, что в металлических сплавах с электронной концентрацией $n \gtrsim 10^{22} \text{ см}^{-3}$ переходы металл–изолятор вообще не наблюдаются (см. гл. 1).

5.5. Формула Ландауэра для одномерных (1D) систем ¹⁾

Модель Андерсона исследована и для систем пониженной размерности. Как это обычно бывает при фазовых переходах, размерность является чрезвычайно существенным параметром. Мы уже сталкивались с этим в гл. 2 при описании слабой локализации. Критерий понижения размерности там был очень мягкий, поскольку характерный размер объекта b сравнивался в формуле (2.8) со сравнительно большой диффузионной длиной L_φ . Здесь речь идет о сильной локализации, и критерий гораздо более жесткий. Он связан со структурой электрон-

¹⁾ Последовательное изложение теоретических аспектов этого вопроса см. в [9].

ного спектра свободных электронов с волновыми функциями $\exp(i\mathbf{k}\mathbf{r})$ при соответствующей геометрии области их существования

$$\varepsilon = \frac{\hbar^2 k_{\parallel}^2}{2m} + \varepsilon_{\perp}(i), \quad i = 1, 2, \dots \quad (5.21)$$

Здесь $k_{\parallel}$ — волновой вектор в направлениях, где движение электронов не ограничено, $\varepsilon_{\perp}$ — размерно квантованная часть энергии, связанная с движением в ограниченных направлениях, а i — номер размерно квантованной подзоны. В пленке $k_{\parallel}^2 = k_x^2 + k_y^2$, а вдоль нормали Oz устанавливается стоячая волна, и энергия движения квантуется. В проволоке $k_{\parallel}$ — это волновой вектор вдоль ее оси, а $\varepsilon_{\perp}(i)$ определяется квантованием в двух поперечных направлениях.

Система имеет пониженную размерность, если все электроны помещаются в нижней размерно квантованной подзоне. Для вырожденной электронной системы критерий имеет вид

$$\varepsilon_F < \Delta_s, \quad \Delta_s \equiv \varepsilon_{\perp}(i=2) - \varepsilon_{\perp}(i=1) \propto b^{-2}. \quad (5.22)$$

Последующая часть этой главы посвящена именно системам приближенной размерности, в основном одномерным.

Соединим два резервуара, к которым приложена разность потенциалов V , идеальной проволокой длины L . «Идеальная» означает, что в проволоке полностью отсутствует рассеяние, даже упругое. Тогда всякий электрон, попадающий в проволоку с одной стороны, с вероятностью единица выходит с другой. Пусть к тому же диаметр проволоки столь мал, что в ее спектре (5.21) под уровень Ферми ε_F попадает ограниченное число $\nu = 2N_s$ размерно квантованных подзон (их также называют каналами; в отсутствие магнитного поля при каждом $i \leq N_s$ имеется два канала с разными направлениями спинов):

$$\varepsilon_{\perp}(i) < \varepsilon_F \quad \text{при} \quad i = 1, 2, \dots, N_s. \quad (5.23)$$

Если $N_s = 1$, то 1D-систему называют одноканальной (с учетом спина ее можно было бы также называть двухканальной), при $N_s > 1$ она называется многоканальной. Ввиду идеальности проволоки каналы внутри нее независимы и не обмениваются электронами. Плотность электронов n_i в канале i , продольная скорость электронов v_i и плотность состояний g_i на уровне Ферми связаны соотношениями

$$v_i = \frac{1}{\hbar} \left(\frac{\partial \varepsilon}{\partial k} \right)_{\varepsilon=\varepsilon_i}, \quad g_i = \left(\frac{\partial n_i}{\partial \varepsilon} \right)_{\varepsilon=\varepsilon_i} = \frac{1}{2\pi\hbar v_i}, \quad (5.24)$$

$$\varepsilon_i = \varepsilon_F - \varepsilon_{\perp}(i), \quad 2 \sum_{i=1}^{N_s} n_i = n.$$

Наличие между резервуарами разности потенциалов V означает, что из-за разности электронной плотности $\delta n_i = g_i eV$ имеется разность потоков электронов, попадающих в канал i справа и слева. В выражении

для тока конкретные параметры канала, фигурирующие в соотношениях (5.24), сокращаются, так что ток в канале J_i не зависит от индекса i и равен

$$J_i = ev_i \delta n_i = \frac{e^2}{2\pi\hbar} V. \quad (5.25)$$

Кондактанс $Y_{id} = J/V$ и сопротивление $R_{id} = 1/Y_{id}$ такой проволоки определяются полным током $J = \sum_1^\nu J_i$ и равны

$$Y_{id} = \frac{e^2}{2\pi\hbar} \nu, \quad R_{id} = \frac{2\pi\hbar}{e^2} \frac{1}{\nu}. \quad (5.26)$$

Индекс в обозначениях подчеркивает, что формула (5.26) относится к идеальной проволоке.

Результат (5.26) замечателен в нескольких отношениях. Во-первых, оказалось, что в 1D-системе, даже в многоканальной, диссипация имеется даже в отсутствие рассеяния. Это проявление принципа нелокальности. Электроны забирают энергию от поля, когда они находятся внутри проволоки, а отдают ее вне проволоки, когда термализуются в резервуаре. Во-вторых, как это ни удивительно, сопротивление проволоки R_{id} не зависит от ее длины и определяется только квантованием электронного спектра.

Казалось бы, утверждение, что сопротивление проволоки не зависит от ее длины, противоречит простому рассуждению: разделим мысленно идеальную проволоку на две части, которые при этом окажутся включенными последовательно; если у каждой части сопротивление R_{id} , то полное сопротивление должно было бы быть $2R_{id}$. Но просто разделить проволоку на две части недостаточно; для того чтобы обе части превратились в независимые сопротивления, между ними нужно вставить дополнительный резервуар-термостат, который бы сделал проходящие через него электронные волны некогерентными. Если температура проволоки отлична от абсолютного нуля, $T \neq 0$, так что существует конечная длина $L_\varphi < \infty$, на которой происходит сбой фазы, то такие термостаты как бы появляются автоматически на расстоянии L_φ друг от друга.

Таким образом температура накладывает на длину идеальной проволоки ограничение сверху, $L < L_\varphi$. Ограничением снизу является ее диаметр, т.е. проволока может быть очень короткой. Это дает возможность проверить формулу (5.26) экспериментально, потому что на коротком участке сравнительно проще добиться отсутствия дефектов.

На рис. 5.11 приведены результаты измерений проводимости узкого канала под расщепленным затвором, соединяющего две области 2D-электронного газа в гетероструктуре GaAs/Al_xGa_{1-x}As [10]. При увеличении запирающего напряжения V_g на затворе обедненная область несколько расширяется за счет того, что она сильнее выступает

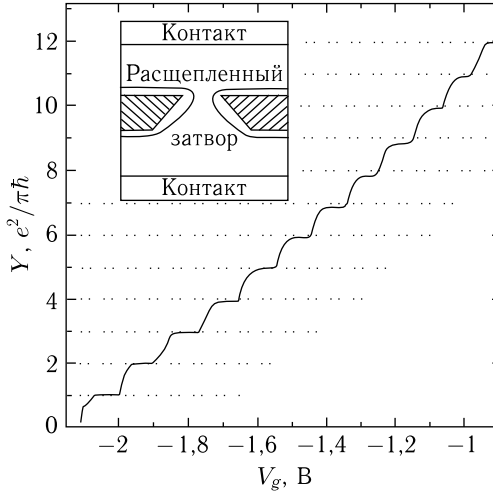


Рис. 5.11. Кондактанс Y баллистического контакта между двумя 2D-областями гетероструктуры GaAs/ $\text{Al}_x\text{Ga}_{1-x}\text{As}$ в зависимости от напряжения на затворе, регулирующем ширину контакта [10]. На вставке схема измерительной ячейки

за края затвора. Как видно из схемы на вставке к рис. 5.11, проводящий канал при этом сужается, что означает уменьшение числа каналов N_s . То, что канал короткий, не мешает применять к нему формулу (5.26), и в то же время именно благодаря этому в нем удастся получить баллистический режим, т.е. отсутствие рассеяния. В структуре, демонстрируемой на рис. 5.11, электронная плотность $3,56 \cdot 10^{11} \text{ см}^{-2}$, длина пробега при 0,6 К около 8,5 мкм, а характерные размеры канала порядка 0,25 мкм.

Из вставки к рис. 5.11 видно, что измерение происходит по двух-контактной схеме, так что в измеряемое сопротивление R_{mea} входит сопротивление R_{cont} контактов и прилегающих к ним широких участков 2D-слоя. Интересующий нас кондактанс Y получается после вычитания R_{cont} из R_{mea} : $Y \equiv R^{-1} = (R_{\text{mea}} - R_{\text{cont}})^{-1}$. В качестве R_{cont} была выбрана величина 4,35 кОм, что примерно соответствует результатам независимых измерений. После вычитания этой величины функция $Y(V_g)$ превращается в последовательность ступеней одинаковой высоты

$$\Delta Y_{\text{id}} = \frac{e^2}{2\pi\hbar} \Delta\nu = \frac{e^2}{\pi\hbar}, \quad (5.27)$$

в полном соответствии с формулой (5.26).

Заметьте. Рассуждения, которые привели к формуле (5.26), не предопределяют распределение электрического поля вдоль проволоки. Например, в краевых каналах, которые образуются при

квантовом эффекте Холла вдоль края образца между контактами и являются идеальными одномерными каналами, все падение напряжения V сосредоточено на границе с одним из контактов (см. рис. 9.15 в гл. 9 и поясняющий его текст).

Откажемся от идеальности проволоки, при этом ограничившись для простоты одноканальной системой. Пусть в заштрихованной части проволоки (рис. 5.12) имеются упругие рассеиватели. Уточнять их взаимное расположение не требуется — будем пока рассматривать всю область как единый рассеивающий объект. В квантовой механике он характеризуется в одномерном случае комплексными коэффициентами отражения r и прохождения t , которые связывают амплитуды отражен-

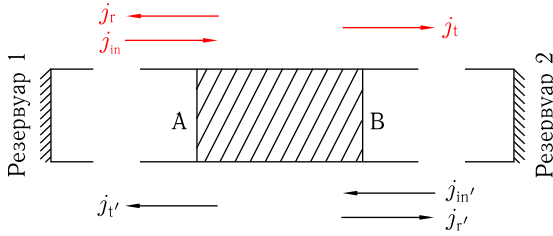


Рис. 5.12. Одномерный проводник, соединяющий два резервуара и состоящий из двух идеальных участков по краям и рассеивающего участка АВ в середине

ной и прошедшей волн с амплитудой падающей волны. Слева и справа на заштрихованную область падают электронные потоки j_{in}/e и $j_{in'}/e$, каждый из которых отражается с вероятностью $\mathcal{R} = |r|^2$ и проходит с вероятностью $\mathcal{T} = |t|^2$. Из симметрии квантовомеханической задачи

$$\mathcal{R} = \frac{j_r}{j_{in}} = \frac{j_{r'}}{j_{in'}}, \quad \mathcal{T} = \frac{j_t}{j_{in}} = \frac{j_{t'}}{j_{in'}}, \quad \mathcal{R} + \mathcal{T} = 1. \quad (5.28)$$

Если падение напряжения на заштрихованной области равно нулю, то и суммарный поток электронов в проволоке равен нулю. При наличии разности потенциалов δV на границах области появляется разность плотностей $\delta n = ge\delta V$. В одномерных системах все электроны движутся вдоль проводника и потому принадлежат к одному из потоков, фигурирующих в уравнениях (5.28). Это позволяет связать δn с плотностями электронов в потоках и выразить δV через токи:

$$\delta V = \frac{\delta n}{ge} = \frac{j_{in} + j_r + j_{t'}}{e^2 gv} - \frac{j_{in'} + j_{r'} + j_t}{e^2 gv} = \frac{2\mathcal{R}(j_{in} - j_{in'})}{e^2 gv}. \quad (5.29)$$

Здесь g и v это плотность состояний и модуль скорости электронов на уровне Ферми. Поскольку полный ток J равен

$$J = j_{in} - j_r - j_{t'} = j_{in'} - j_{r'} - j_t = \mathcal{T}(j_{in} - j_{in'}), \quad (5.30)$$

отношение $J/\delta V$ позволяет написать кондактанс $Y_{\text{imp}} = J/\delta V$ и сопротивление $R_{\text{imp}} = Y_{\text{imp}}^{-1}$ заштрихованной области

$$Y_{\text{imp}} = \frac{e^2}{2\pi\hbar} \frac{\mathcal{T}}{\mathcal{R}} = \frac{e^2}{2\pi\hbar} \frac{\mathcal{T}}{1-\mathcal{T}}, \quad R_{\text{imp}} = \frac{2\pi\hbar}{e^2} \frac{\mathcal{R}}{\mathcal{T}} = \frac{2\pi\hbar}{e^2} \frac{\mathcal{R}}{1-\mathcal{R}}. \quad (5.31)$$

Идея представлять упруго рассеивающие центры в виде потенциальных барьеров на пути распространяющихся волн и выражать транспортные характеристики системы через коэффициенты отражения и прохождения волны через эти барьеры принадлежит Ландауэру. Поэтому соответствующие формулы, в частности выражение для кондактанса (5.31), называют его именем. В принципе техника Ландауэра применима к системам любой размерности, но она особенно удобна и часто используется для 1D-систем.

Заметьте. Формула Ландауэра в виде (5.31) выведена в предположении, что разность потенциалов приложена непосредственно к рассеивающей области между точками А и В на рис. 5.12. Именно поэтому кондактанс (5.31) при слабом рассеянии, $\mathcal{T} \sim 1$, $\mathcal{R} \ll 1$, может оказаться больше, чем кондактанс (5.26) системы, вообще не имеющей рассеивателей.

Если разность потенциалов в системе на рис. 5.12 приложена к резервуарам, то сопротивления идеальной проволоки и области рассеяния включены последовательно и кондактанс Y_{total} всей системы равен

$$Y_{\text{total}}^{-1} = Y_{\text{id}}^{-1} + Y_{\text{imp}}^{-1} \equiv R_{\text{id}} + R_{\text{imp}} = \frac{2\pi\hbar}{e^2} \left(1 + \frac{\mathcal{R}}{\mathcal{T}}\right), \quad (5.32)$$

$$Y_{\text{total}} = \frac{e^2}{2\pi\hbar} \mathcal{T}.$$

Теперь при $\mathcal{T} \rightarrow 1$ кондактанс $Y \rightarrow Y_{\text{id}}$, как и должно быть.

Выражение (5.32) для Y_{total} можно получить и непосредственно, приложив разность потенциалов к резервуарам, записав электронный поток из одного резервуара в другой и учтя однократное рассеяние (ср. с выводом формулы (5.26)). Это означает, что сложение сопротивлений в соответствии с законом Ома в рассуждениях (5.32) было правомерным. Как мы покажем в следующем параграфе, в одномерных системах из-за интерференции падающих и отраженных волн это отнюдь не всегда так.

5.6. Локализация и роль корреляций в 1D-системах

Рассмотрим два последовательных барьера в одноканальном одномерном проводнике (рис. 5.13) и выразим параметры $\mathcal{T}_s$ и $\mathcal{R}_s = 1 - \mathcal{T}_s$ образовавшегося составного рассеивающего объекта через параметры $\mathcal{T}_1$, $\mathcal{R}_1$, $\mathcal{T}_2$ и $\mathcal{R}_2$ исходных барьеров. Если на барьер 1 слева падает волна амплитуды I , то сформировавшееся стационарное волновое поле

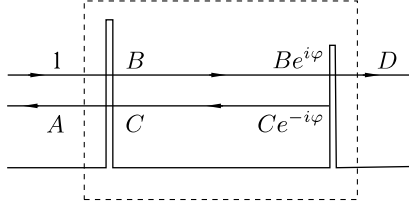


Рис. 5.13. Рассеивающий участок в 1D-проводнике, состоящий из двух барьеров. Комплексные амплитуды $A, \dots, D$ волн, приходящих и уходящих от обоих барьеров, все нормированы на амплитуду исходной приходящей волны, помеченной единицей

будет содержать еще четыре волны: отраженную A , прошедшую D и две волны между барьерами, B и C , движущиеся в противоположные стороны ($A, \dots, D$ — это комплексные амплитуды волн). Выразив амплитуды волн, уходящих направо и налево от каждого из барьеров, через амплитуды падающих волн, получим четыре уравнения

$$A = r_1 + Ct_1, \quad B = t_1 + Cr_1, \quad Ce^{-i\varphi} = Be^{i\varphi}r_2, \quad D = Be^{i\varphi}t_2. \quad (5.33)$$

Здесь использовано, что коэффициенты отражения от барьера не зависят от того, с какой стороны падает волна, $r_1 = r'_1$; множители $\exp(\pm i\varphi)$ учитывают набег фазы волны на расстоянии от одного барьера до другого. Из уравнений (5.33) следует

$$D = \frac{e^{i\varphi}t_1t_2}{1 - e^{2i\varphi}r_1r_2}, \quad T_s = |D|^2 = \frac{T_1T_2}{1 + \mathcal{R}_1\mathcal{R}_2 - 2\sqrt{\mathcal{R}_1\mathcal{R}_2} \cos \vartheta}, \quad (5.34)$$

где $\vartheta = 2\varphi + \arg(r_1r_2)$. Кондактанс Y_s составного «двухбарьерного» рассеивателя, который выделен на рис. 5.13 пунктиром, равен

$$Y_s = \frac{e^2}{2\pi\hbar} \frac{T}{1 - T} = \frac{e^2}{2\pi\hbar} \frac{T_1T_2}{\mathcal{R}_1 + \mathcal{R}_2 - 2\sqrt{\mathcal{R}_1\mathcal{R}_2} \cos \vartheta}. \quad (5.35)$$

Если составной «двухбарьерный» рассеиватель состоит из двух одинаковых барьеров, $r_1 = r_2 = r$, $t_1 = t_2 = t$, $\mathcal{R}_1 = \mathcal{R}_2 = \mathcal{R}$ и т. д., то

$$Y_s = \frac{e^2}{2\pi\hbar} \frac{T^2}{4\mathcal{R} \sin^2 \vartheta/2}, \quad \vartheta/2 = \varphi + \arg(r) = k\tilde{l} + \arg(r), \quad (5.36)$$

где k — волновой вектор, а $\tilde{l}$ — расстояние между барьерами.

Кондактанс (5.35) зависит не только от параметров двух исходных барьеров; через угол ϑ он зависит и от расстояния между ними. Поскольку нас в конечном счете интересует 1D-проводник с большим количеством случайно расположенных барьеров, то можно усреднить по всем возможным расстояниям между ними, предположив, что угол ϑ с одинаковой вероятностью принимает любые значения от 0 до 2π . Такое усреднение не совсем корректно, но оно позволяет нам продвинуться дальше и проследить тенденции, возникающие при удлинении

нии цепочки одномерных барьеров (более подробно см. [9], а также оригинальную работу [11]). Из среднего значения $\cos \vartheta = 0$ следует усредненный кондактанс $\bar{Y}_s$ системы из двух барьеров. Он равен

$$\bar{Y}_s = \frac{e^2}{2\pi\hbar} \frac{T_1 T_2}{\mathcal{R}_1 + \mathcal{R}_2}. \quad (5.37)$$

Для сравнения выпишем классическое выражение $Y_s^{(cl)}$ для кондактанса двух последовательных сопротивлений R_1 и R_2

$$\begin{aligned} Y_s^{(cl)} &= (R_1 + R_2)^{-1} = \frac{e^2}{2\pi\hbar} \left(\frac{\mathcal{R}_1}{1 - \mathcal{R}_1} + \frac{\mathcal{R}_2}{1 - \mathcal{R}_2} \right)^{-1} = \\ &= \frac{e^2}{2\pi\hbar} \frac{T_1 T_2}{\mathcal{R}_1 + \mathcal{R}_2 - 2\mathcal{R}_1 \mathcal{R}_2}. \end{aligned} \quad (5.38)$$

В нем имеется лишний по сравнению с (5.37) член в знаменателе, пропорциональный произведению коэффициентов пропуска двух барьеров $\mathcal{R}_1 \mathcal{R}_2$. Разница между выражениями (5.37) и (5.38) становится особенно наглядной, если написать их для безразмерных сопротивлений $\varrho_i = \mathcal{R}_i / T_i$ ($i = 1, 2, s$):

$$\begin{aligned} \bar{\varrho}_s &\equiv \frac{e^2}{2\pi\hbar} (\bar{Y}_s)^{-1} = \frac{\mathcal{R}_1}{T_1} + \frac{\mathcal{R}_2}{T_2} + 2 \frac{\mathcal{R}_1}{T_1} \frac{\mathcal{R}_2}{T_2} = \varrho_1 + \varrho_2 + 2\varrho_1 \varrho_2, \\ \varrho_s^{(cl)} &= \varrho_1 + \varrho_2. \end{aligned} \quad (5.39)$$

Рассмотрим длинную цепочку одинаковых, но расположенных на случайных расстояниях $\tilde{l}_i$ друг от друга слабо рассеивающих барьеров $\mathcal{R} \ll 1$, $T \sim 1$, имеющих каждый малое сопротивление

$$R = (2\pi\hbar/e^2)(\mathcal{R}/T) \approx (2\pi\hbar/e^2)\mathcal{R} \ll 2\pi\hbar/e^2.$$

Будем вычислять сопротивление

$$R_N = Y_N^{-1} = (2\pi\hbar/e^2)(\mathcal{R}_N/T_N)$$

составного рассеивающего объекта из N барьеров по рекуррентной формуле, следующей из (5.37),

$$\frac{\mathcal{R}_N}{T_N} = \frac{\mathcal{R}_{N-1} + \mathcal{R}}{T_{N-1} T}. \quad (5.40)$$

Пока число N барьеров мало, а именно, пока $N\mathcal{R} \ll 1$, величина T_N остается порядка единицы, а величина $\mathcal{R}_N \approx N\mathcal{R}$. В этих условиях сопротивление R_N растет практически линейно: $R_N \approx NR \propto N$. Из сравнения $1/R_N$ с формулой для удельной проводимости $\sigma = (ne^2 l)/(\hbar k_F)$, в которую в качестве n подставлена концентрация электронов в одном канале (в частности, с одним направлением спина), летящих по полю, $n = n^* = k_F/2\pi$, получим для эффективной длины свободного пробега l в проволоке длиной L выражение

$$l = \bar{l}/\mathcal{R} \quad (L = N\bar{l} \ll \bar{l}/\mathcal{R}), \quad (5.41)$$

где $\bar{l} = \sum \tilde{l}_i / N$ — среднее расстояние между барьерами. Отсюда сопротивление «одноканальной проволоки» равно

$$R(L) = \frac{2\pi\hbar}{e^2} (L/\bar{l}) \quad (L \ll \bar{l}/\mathcal{R}). \quad (5.42)$$

Поскольку $\mathcal{R}_N$ не может стать больше единицы, можно в формуле (5.40) начиная с некоторого $N \gg \mathcal{R}^{-1}$, положить $\mathcal{R}_N \approx \mathcal{R}_{N-1} \approx 1$. Отсюда сразу следует, что при $N \rightarrow \infty$

$$\begin{aligned} T_N &\approx T_{N-1}T, & T_N &\rightarrow sT^N = se^{aN} \\ (s = \text{const}, & a = \ln T < 0). \end{aligned} \quad (5.43)$$

Экспоненциальное уменьшение интенсивности прошедшей волны T_N при росте числа N , т.е. при росте длины проволоки $L = \bar{l}N$, является демонстрацией 1D-локализации на конкретном примере. Используя понятие радиуса локализации (4.1) и выражение (5.32) для кондактанса Y_{total} проволоки, соединяющей два термостата, формулу (5.43) можно переписать в виде

$$Y_{\text{total}}(L) = \frac{e^2}{2\pi\hbar} T_N \approx \frac{e^2}{2\pi\hbar} sT^N \approx \frac{e^2}{2\pi\hbar} se^{-L/\xi}, \quad \xi = \bar{l}/|\ln T|. \quad (5.44)$$

Поскольку $\mathcal{R} \ll 1$, то $|\ln T| = |\ln(1 - \mathcal{R})| \approx \mathcal{R}$, так что выражения (5.41) для l и (5.44) для ξ совпадают: длина свободного пробега, описывающая металлическую проводимость при сравнительно небольших N в интервале $1 \ll N \ll \mathcal{R}^{-1}$, при достаточно больших N превращается в длину локализации ξ ,

$$l = \xi \quad \text{при} \quad \nu = 1. \quad (5.45)$$

Остановимся на еще одной особенности транспорта в 1D-системах. На рис. 5.14 приведены транспортные характеристики квази 1D-системы, изготовленной на базе аккумулирующего слоя в полевом транзисторе на поверхности n-Si [12]. При низкой температуре на зависимости кондактанса Y от напряжения на затворе V_g появляется шумоподобная составляющая очень большой амплитуды. Это не настоящий шум. Сигнал не зависит от времени и, если не отогревать образец до комнатной температуры, то при повторном эксперименте кривая $Y(V_g)$ воспроизводится вплоть до мельчайших подробностей. Видно, что при низких температурах и при напряжениях на затворе V_g , обеспечивающих узкий канал и малую концентрацию носителей, кондактанс испытывает хаотические узкие осцилляции при изменении V_g , размах которых растет при понижении температуры. На другом образце, и даже на этом же при повторном охлаждении от комнатной температуры, детальная структура осцилляций другая при той же общей картине эволюции осцилляций с изменением температуры и напряжения V_g .

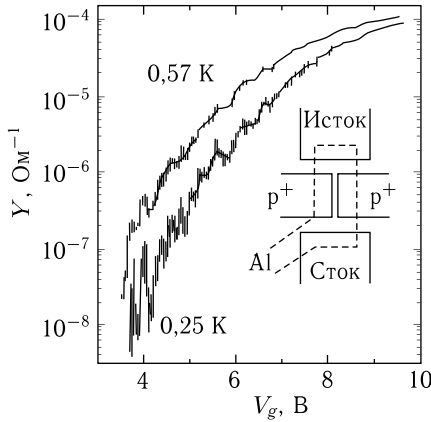


Рис. 5.14. Зависимость от напряжения на затворе V_g кондактанса Y длинного квазиодномерного канала в полевом транзисторе, изготовленном на поверхности p-Si, и находящемся в режиме аккумулирующего слоя [12]. Ширина канала может меняться от нуля до максимума ~ 1 мкм, заданного конструкцией (см. схему на вставке), при помощи напряжений на контрольных электродах p^+ и на затворе

Фундаментальная причина хаотических осцилляций в одномерности. Все дефекты в проволоке включены последовательно и линии тока не могут обойти ни один из них. Выключение одного дефекта, осуществляющего сильное рассеяние, может поэтому сильно повлиять на суммарное сопротивление. Вопрос в том, как изменение V_g , которое меняет концентрацию носителей и их энергию Ферми ε_F , может включать, или выключать, или менять эффективность отдельных дефектов.

Вернемся к выражению (5.36) для кондактанса Y_s симметричного «двухбарьерного» рассеивателя. Выше мы усредняли это выражение по $\cos \vartheta$ на том основании, что имеется разброс расстояний $\tilde{l}_i$ между барьерами. Но входящий в ϑ угол $\varphi = k\tilde{l}_i$ зависит не только от $\tilde{l}_i$, но и от волнового вектора k , т. е. от энергии рассеивающегося электрона ε_F . Для одной конкретной рассеивающей пары барьеров с фиксированным значением $\tilde{l}_i$ из формулы (5.36) следует, что R_s принимает значение в интервале от нуля до $4R$,

$$0 \leq R_s \leq 4R, \quad (5.46)$$

в зависимости от энергии налетающего электрона.

Уместно напомнить две вещи:

— транспортные свойства 1D-системы определяются именно электронами из окрестности ε_F , потому что противоположные потоки электронов с меньшими энергиями компенсируют друг друга.

— выше мы упоминали о существовании проблем с усреднением выражения для сопротивления (5.35); они связаны именно с большим диапазоном (5.46) изменения величины R_s .

Пространство между двумя барьерами представляет собой потенциальную яму. В этой яме вообще говоря имеется набор уровней ε_i , ширина которых обусловлена прозрачностью барьеров t_1 и t_2 . По мере того как энергия электрона ε_F смещается относительно системы уровней в этой яме, вероятность туннелирования осциллирует, достигая максимума в условиях резонанса $\varepsilon_F = \varepsilon_i$. Поэтому гигантские хаотические осцилляции сопротивления на рис. 5.14 можно теоретически описать именно в терминах резонансного туннелирования.

Модель локализованных состояний в 1D-системах использует представления об электронных уровнях внутри составных рассеивателей. При достаточно низких температурах отражения от далеких барьеров

$$1 \ll N \ll \frac{L_\varphi}{l} \quad (5.47)$$

остаются когерентными. Поэтому, согласно соотношениям (5.43), эти отражения при достаточно большом L_φ скомпенсируют прозрачность барьеров t_1 и t_2 и сделают состояние между ними истинно локализованным. В этих условиях следует ожидать прыжковый характер проводимости. И действительно, на рис. 5.15 приведены измерения темпера-

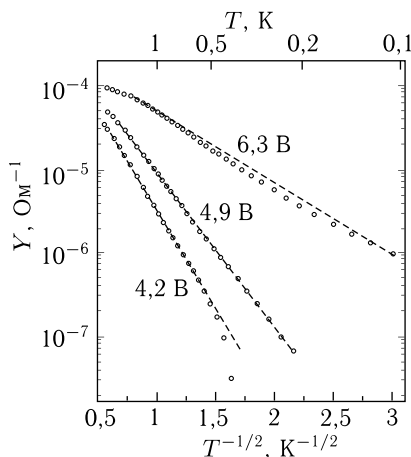


Рис. 5.15. Графики температурной зависимости в минимумах кондактанса канала полевого транзистора при трех разных значениях напряжения на затворе V_g в осях $(T^{-1/2}, \log Y)$ [12]

турной зависимости кондактанса, сделанные в нескольких минимумах кривой, приведенной на рис. 5.14. Видно, что при измерениях в левой части графика на рис. 5.14, при меньших V_g , когда кондактанс мал,

осцилляции велики и есть все основания считать канал одномерным, точки хорошо ложатся на функциональную зависимость

$$Y = Y_0 \exp[-(T_M/T)^{1/2}], \quad (5.48)$$

в полном соответствии с формулой Мотта (4.24). При больших V_g канал расширяется и постепенно превращается в двумерный. Кондактанс при этом увеличивается, а амплитуда хаотических осцилляций падает. Экспериментальные точки зависимости $\log Y(T^{-1/2})$, снятые при напряжении на затворе $V_g = 6,3$ В, отклоняются на рис. 5.15 от прямой, но спрямляются в осях $(\log Y, T^{-1/3})$, опять-таки в полном соответствии с формулой (4.24).

Вернемся к общему утверждению о неминуемой локализации на 1D случайном потенциале. Уже в формуле (5.36), на базе которой была построена конкретная модель (5.43), демонстрирующая локализацию, содержится намек на то, что можно попытаться избежать локализации за счет корреляций в случайном потенциале. Из этой формулы следует, что для «двухбарьерного рассеивателя» существует волновой вектор $k_0 = -\arg(r)/\tilde{l}$, при котором этот рассеиватель полностью прозрачен для падающей волны, и отраженной волны нет, $\mathcal{R}_s = 0$. Если заменить в нашей модели (5.40)–(5.43) одиночные барьеры на сдвоенные (5.36), то электрон с энергией $\varepsilon_0 = \hbar^2 k_0^2 / 2m$ окажется делокализованным.

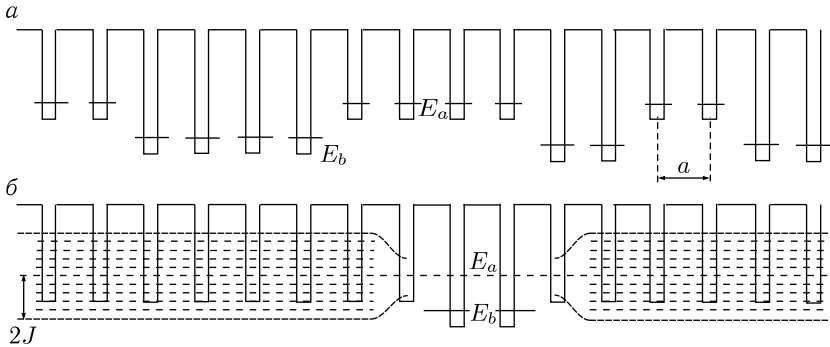


Рис. 5.16. а) Димерная модель одномерного случайного потенциала. Изменения положений уровней из-за перекрытия ям не показано. б) Электронные уровни в одномерной решетке с одним димерным дефектом

Эта идея была развита более подробно в так называемой димерной модели [13]. В ней используются не случайно расположенные барьеры, а одномерная цепочка периодически расположенных потенциальных ям. Цепочка состоит из ям двух сортов, с уровнями энергии E_a и E_b . При этом ямы распределены по нечетным узлам решетки совершенно случайно, без каких бы то ни было корреляций, а в каждом четном узле находится яма того же сорта, что и нечетном узле слева от него. Это означает, что одинаковые ямы стоят парами, откуда и название

модели (рис. 5.16а). Если расстояние между ямами a , то получившуюся решетку можно представить как сумму двух случайных, но *одинаковых* подрешеток, сдвинутых на a друг относительно друга, обе с периодом $2a$ и с совершенно случайным распределением ям по узлам.

Будем считать пары с энергией E_a принадлежащими основной решетке, а пары с энергией E_b — дефектами. Как мы уже видели, в этой модели могут существовать делокализованные состояния при некоторых выделенных значениях энергии, обусловленных структурой дефекта. Условие того, что электрон с такой энергией может распространяться в основной решетке, можно сформулировать при помощи рис. 5.16 б, где один димерный дефект из двух ям с энергиями E_b помещен в идеальную решетку из ям E_a . Пусть интеграл перекрытия между соседними ямами равен J . Тогда справа и слева от дефекта образуются зоны с квазинепрерывным распределением уровней $\varepsilon = E_a - 2J \cos ka$. Если выполняется соотношение

$$|E_a - E_b| < 2J, \quad (5.49)$$

то невозмущенный энергетический уровень дефекта E_b попадает внутрь зоны, и в зоне появляется выделенное значение $k = k_0$, $\cos k_0 a = (E_a - E_b)/2J$, для которого вероятность отражения от дефекта $\mathcal{R} = 0$.

В димерной модели корреляции существуют только между ближайшими соседями. При таких корреляциях делокализованные состояния возникают только при дискретных значениях энергии. Для того чтобы получить полосу делокализованных состояний, необходимо использовать дальние корреляции, сохранив при этом в потенциале элемент случайности. Алгоритм построения такого потенциала был предложен в работе [14]. Мы приведем в следующем параграфе конкретный пример такого алгоритма, составленного для экспериментальной проверки этих идей путем микроволнового моделирования.

5.7. Микроволновое моделирование ¹⁾

Зависящее от времени уравнение Шредингера

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi + U \Psi \quad (5.50)$$

и классическое волновое уравнение

$$\frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} = \Delta \Psi - U \Psi \quad (5.51)$$

¹⁾ Более подробно эти вопросы излагаются в книге [15].

(c это скорость волны), после подстановки $\Psi = e^{-i\omega t}\psi$ сводятся к одному и тому же уравнению

$$(\Delta - U + k^2)\psi = 0 \quad (5.52)$$

с той лишь разницей, что для уравнения Шредингера

$$\omega = \frac{\hbar}{2m} k^2, \quad (5.53)$$

а для волнового уравнения

$$\omega = ck. \quad (5.54)$$

Это дает возможность моделировать локализационные процессы в устройствах с высокочастотными электромагнитными полями. Мы приведем два примера такого моделирования, для одномерной и для двумерной систем.

В работе [16] одномерный случайный потенциал имитировали длинным волноводом с рассеивателями внутри, а измеряли коэффициент прохождения электромагнитной волны в зависимости от частоты. Схема волновода изображена на рис. 5.17. Рабочий диапазон частот ν был выбран внутри частотного диапазона, в котором волновод находился в одномодовом режиме: $7,5 \text{ ГГц} = c/2a < \nu < c/a = 15 \text{ ГГц}$, где $a = 20 \text{ мм}$ это больший размер сечения волновода.

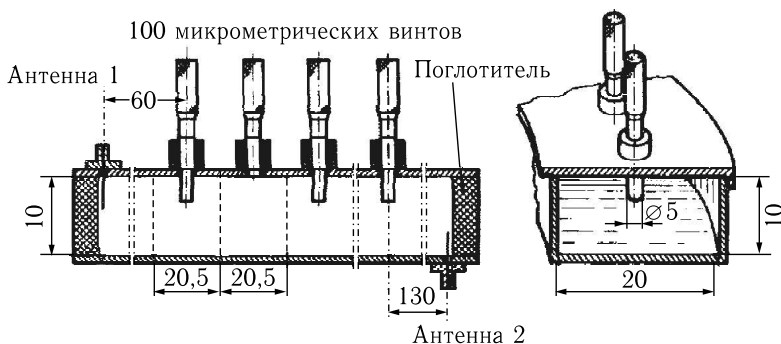


Рис. 5.17. Схематический чертеж одномодового волновода со 100 рассеивателями, в котором измерялся коэффициент прохождения t электромагнитной волны в зависимости от частоты [16]. Все размеры в миллиметрах

На равных расстояниях по длине волновода в него введены $N = 100$ штырьков-рассеивателей, моделирующих случайный потенциал. Они могут при помощи микрометрических винтов вдвигаться на разную глубину u_n , где $1 \leq n \leq N$. Глубина устанавливается по формуле

$$u_n = \sqrt{u_n^2} \sum_{m=-\infty}^{\infty} \beta_m Z_{n+m}, \quad \beta_m = \frac{2}{\pi} \int_0^{\pi/2} \sqrt{\varphi(\mu)} \cos(2m\mu) d\mu. \quad (5.55)$$

Здесь Z_{n+m} случайные числа в интервале от -1 до $+1$. Именно они вносят в потенциал элемент случайности. Корреляции между всеми u_n обеспечиваются множителями β_m , определенными через функцию $\varphi(\mu)$. Последняя выбирается при помощи специального математического алгоритма в зависимости от того, какой требуется спектр пропускания одномерной системы. Пример реализации такой программы показан на рис. 5.18. В соответствии с имеющимся алгоритмом была выбрана функция $\varphi(\mu)$, при которой должны были возникнуть

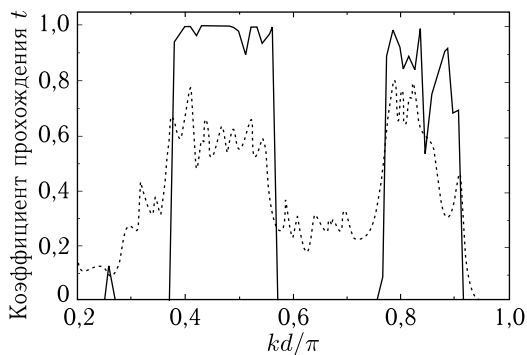


Рис. 5.18. Коэффициент прохождения t волны в одномерном канале при наличии периодически расположенных случайных рассеивателей, между которыми имеются корреляции, со специально выбранной корреляционной функцией. Сплошная линия — численный эксперимент с $N = 10^4$ рассеивателями; пунктир — микроволновый эксперимент с $N = 100$ рассеивателями, усредненный по пяти разным реализациям. Корреляционная функция одна и та же [16]

две полосы пропускания внутри рабочего диапазона. Сплошной линией на рисунке показан определенный в численном эксперименте коэффициент пропускания волновода с одномерной последовательностью из $N = 10^4$ рассеивателей, выбранных в соответствии с формулой (5.55) по заданной функции $\varphi(\mu)$. Пунктиром показан усредненный по пяти разным реализациям результат реального микроволнового эксперимента с последовательностью из $N = 100$ рассеивателей.

Наличие полосы пропускания конечной ширины означает на языке физики твердого тела существование края подвижности, а следовательно, при выборе подходящего управляющего параметра, и перехода металл–изолятор. Таким образом на примере одномерных моделей, где удастся продвинуться достаточно далеко и аналитически, и при помощи численных методов, продемонстрирована особая роль корреляций случайного потенциала в вопросах локализации.

Микроволновую технику можно использовать и для моделирования задач по двумерной локализации. Соответствующие эксперименты называются микроволновыми бильярдами. Их основными элементами являются плоские резонаторы, запитываемые через подводящую антенну.

Поперечный размер такого резонатора вдоль оси z меньше или порядка длины волны, а продольные размеры вдоль осей x и y много больше длины волны. Случайный потенциал U моделируется металлическими рассеивателями, хаотически расположенными внутри резонатора. Стационарное распределение электромагнитного поля внутри резонатора, например электрического поля $E_z(x, y)$ при соответствующей моде волны, удовлетворяет не только уравнениям Максвелла, но и уравнению Шредингера (5.50), (5.52) для собственной волновой функции ψ . Это распределение можно измерить, например, перемещая внутри резонатора маленький металлический шарик. Возмущение, вносимое таким шариком, приводит к изменению частоты резонатора $\Delta\omega$, пропорциональному квадрату электрического поля в данной точке, $\Delta\omega \propto E_z^2(x, y)$, и, соответственно, квадрату собственной волновой функции электрона, $\Delta\omega \propto \psi^2(x, y)$, при значении энергии $\hbar\omega$.

Для примера на рис. 5.19, взятом из работы [17], приведены три собственные волновые функции в прямоугольном бильярде со случайными рассеивателями, полученные на трех разных частотах. Амплитуда $\psi^2(x, y)$ изображена при помощи разной степени почернения. Этот пример демонстрирует делокализацию электрона при повышении его энергии.

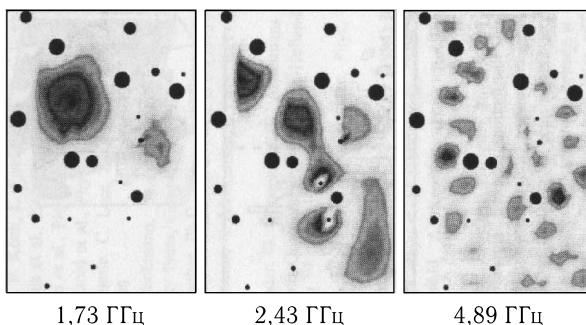


Рис. 5.19. Собственные волновые функции, полученные при разных частотах в прямоугольном бильярде размерами $240 \times 340 \text{ мм}^2$ со случайными рассеивателями (черные кружки) [17]

5.8. Контрольные вопросы и задачи

1. Идеальная одномерная одноканальная проволока имеет конечное сопротивление $\hbar/e^2$. Сопротивление идеальной двух- и трехмерной среды равно нулю. Как происходит переход к нулевому сопротивлению при постепенном увеличении поперечного сечения проволоки?

$$R_\nu = (1/\nu)(\hbar/e^2).$$

2. Сконструировать из четырех одинаковых дельта-образных барьеров составной дефект, который не отражает электроны с энергиями ε_1 и ε_2 .

3. Возможно ли, чтобы составной дефект, состоящий из двух разных барьеров с коэффициентами отражения $r_1 \neq r_2$, имел нулевое отражение на какой-либо определенной энергии?

Нет.

4. Между двумя проводящими полупространствами вставлена тонкая изолирующая плоскость с отверстием диаметром d . Вычислить сопротивление пространства R в направлении, перпендикулярном изолирующей плоскости, если

а) в обоих полупространствах длина пробега электронов $l \ll d$ (контакт Максвелла);

б) в обоих полупространствах длина пробега электронов $l \gg d$ (контакт Шарвина);

в) два полупространства заменить двумя полуплоскостями, а круглое отверстие — щелью, сохранив в этом двумерном случае условие Шарвина $l \gg d$;

г) результаты б) и в) сравнить с формулой (5.26).

С точностью до численных множителей

а) $R = \rho/d$ (ρ — удельное сопротивление);

б) $R = (\rho/d)(l/d) = (h/e^2)(k_F d)^2$;

в) $R = (h/e^2)(k_F d)$.

5. Рассчитать плотность состояний с малыми сдвигами энергии в модели структурного беспорядка.

Шкловский и Эфрос, с. 67; И. М. Лифшиц, УФН 83, 617 (1964).

6. Ионный кристалл соляной кислоты HCl представляет собой совокупность протонов в периодической решетке ионов Cl^- . Почему протоны локализованы, а не образуют зону проводимости?

7. С одной стороны, теорема Блоха утверждает, что в периодическом кристаллическом поле электронные волновые функции всегда делокализованы и имеют вид плоских блоховских волн. С другой стороны, даже в идеально периодической решетке существует переход Мотта, т. е. можно реализовать условия, когда электроны локализованы. Что позволяет вывести эти электроны из-под юрисдикции теоремы Блоха?

Межэлектронное взаимодействие.

Список литературы

1. Mott N. V. Metal-insulator transitions. — Taylor & Francis, 1990.
2. Lee P. A., Ramakrishnan T. V. Disordered electronic systems // Rev. Mod. Phys. **57**, 287 (1985).

3. *Kramer B., MacKinnon A.* // Rep. Prog. Phys. **56**, 1469 (1993).
4. *Шкловский Б. И., Эфрос А. Л.* Электронные свойства легированных полупроводников, § 9. — М.: Наука, 1979.
5. *Лифшиц И. М.* // УФН **83**, 617 (1964).
6. *Edwards P. P., Sienko M. J.* // Phys. Rev. B **17**, 2575 (1978).
7. *Paalanen M. A., Rosenbaum T. F., Thomas G. A., Bhatt R. N.* // Phys. Rev. Lett. **48**, 1284 (1982).
8. *Endo H., Eatah A. I., Wright J. G., and Cusack N. E.* // J. Phys. Soc. Jap. **34**, 666 (1973).
9. *Imry Y.* Introduction to Mesoscopic Physics. — Oxford University Press, 1997 [Есть русский перевод: *Имри Й.* Введение в мезоскопическую физику. — М.: Физматлит, 2002].
10. *van Wees B. J., Kouwenhoven L. P. et al.* // Phys. Rev. B **38**, 3625 (1988).
11. *Anderson P. W., Thouless D. J., Abrahams E., Fisher D. S.* // Phys. Rev. B **22**, 3519 (1980).
12. *Fowler A. B., Harstein A., Webb R. A.* // Phys. Rev. Lett. **48**, 196 (1982).
13. *Dunlap D. H., Wu H.-L., Phillips P. W.* // Phys. Rev. Lett. **65**, 88 (1990).
14. *Izrailev F. M., Krokhin A. A.* // Phys. Rev. Lett. **82**, 4062 (1999).
15. *Stöckmann H.-J.* Quantum Chaos. — Cambridge Univ. Press, 1999 [Есть русский перевод: *Штокман Х.-Ю.* Квантовый хаос (введение). — М.: Физматлит, 2004].
16. *Kuhl U., Izrailev F. M., Krokhin A. A., Stöckmann H.-J.* // Appl. Phys. Lett. **77**, 633 (2000).
17. *Stöckmann H.-J., Barth M., Dörr U., Kuhl U., Schanze H.* // Physica E **9**, 571 (2001).

Глава 6

СКЕЙЛИНГОВАЯ ГИПОТЕЗА ¹⁾

В этой главе обсуждается феноменологическая теория, призванная разрешить противоречие между простейшими соображениями о минимальной металлической проводимости и экспериментами. Она основана на скейлинговой гипотезе. Это название гипотезы происходит от английского слова *scale* — шкала. В русской литературе она часто называется гипотезой подобия.

6.1. Обоснование и формулировка скейлинговой гипотезы ²⁾

Переход металл–изолятор — не совсем обычный фазовый переход. Основное свойство, по которому различаются состояния — проводимость — проявляется лишь в неравновесных условиях. Поэтому одним из оправданий отношения к этому переходу как к непрерывному фазовому переходу может служить успешное описание его при помощи стандартного математического аппарата теории фазовых переходов. Введем управляющий параметр x таким образом, чтобы разность $x - x_c$ имела разные знаки по разные стороны от перехода. Как мы уже знаем из предыдущей главы, таким управляющим параметром может быть электронная концентрация n , уровень беспорядка, магнитное поле B , давление и т. п. Однако это не может быть температура, потому что переход происходит при $T = 0$. Нам нужно также выбрать одну из функций, описывающих транспорт, на роль функции, задающей состояние системы, как это принято в термодинамике.

Пусть наши образцы имеют вид гиперкуба объемом L^d в пространстве размерности d . Проводимость конкретного образца, имеющую размерность $[\text{Ом}^{-1}]$ при любом d , будем называть кондактансом Y . Введем также безразмерный кондактанс y при помощи формулы $(e^2/\hbar)y = Y$, в которой комбинация атомных констант $e^2/\hbar$ имеет размерность $[\text{Ом}^{-1}]$ и равна $2,43 \cdot 10^{-4} \text{Ом}^{-1}$. Величину $e^2/\hbar$ часто

¹⁾ Материал этой главы, как и предыдущей, полезно прочесть и в ином изложении, например, в книге [1] и в обзорах [2, 3].

²⁾ Все основные идеи этого параграфа исчерпывающе изложены в короткой пионерской работе [4].

называют квантом проводимости. Кроме кондактанса существует еще и удельная проводимость с размерностью $[\text{Ом}^{-1}\text{см}^{2-d}]$; будем обозначать ее буквой σ и называть для краткости просто проводимостью. Размерный и безразмерный кондактансы Y и y связаны с проводимостью σ соотношениями

$$Y = \sigma L^{d-2}, \quad y = (e^2/\hbar)^{-1} \sigma L^{d-2}. \quad (6.1)$$

На первый взгляд кажется естественным выбрать в качестве базовой функции состояния проводимость σ . Однако, она лишь до тех пор остается «удельной», т. е. работает как характеристика материала, пока размер образца L остается больше некоторого внутреннего параметра материала размерности длины.

Вспомните: В классической физике металлов, где таким параметром является длина свободного пробега l , существует размерный эффект на постоянном токе: соотношение $Y = \sigma L$, связывающее кондактанс и проводимость при $d = 3$, справедливо лишь до тех пор, пока $L > l$.

Одной из основных величин в теории фазовых переходов является корреляционная длина ξ , которая определена с обеих сторон перехода. На самом переходе, при приближении управляющего параметра x к критическому значению $x \rightarrow x_c$, корреляционная длина ξ расходится: $\xi \rightarrow \infty$. В теории обычных непрерывных фазовых переходов предполагается, что при приближении к переходу ξ — это характерные размеры флуктуаций. Переход металл-изолятор происходит при абсолютном нуле температуры, и вместо тепловых флуктуаций в нем фигурируют квантовые флуктуации. В непосредственной окрестности перехода в течение короткого времени δt могут существовать области с характерными размерами ξ , в которых электронная система находится в энергетически невыгодном основном состоянии. При этом закон сохранения энергии выполняется с точностью до соотношения неопределенности $\delta E \delta t \sim \hbar$.

Поскольку корреляционная длина ξ расходится на переходе, в окрестности x_c условие $L > \xi$ обязательно нарушится, функция σ потеряет смысл, а связь кондактанса и проводимости уже не будет сводиться к множителю L^{d-2} . Это определяет выбор не проводимости, а кондактанса y в качестве базовой функции.

В пользу кондактанса имеются также серьезные аргументы, основанные на анализе волновых функций. Рассмотрим трансформацию дискретного электронного спектра при объединении 2^d гиперкубов объемом L^d в один объемом $(2L)^d$ (эти рассуждения принадлежат Таулессу). Волновые функции в большом кубе ψ_{2L} можно представить в виде линейных комбинаций волновых функций ψ_L в кубах L^d . В первом порядке теории возмущений коэффициенты c_i при поправках к волновой функции имеют вид $J/(\Delta_i E)$, где J — интеграл перекрытия, а $\Delta_i E$ — разность энергий невозмущенных состояний (см.,

например, (5.4)). Величина $\Delta_i E$ порядка расстояния между уровнями размерного квантования, $\Delta_i E \sim (g_d L^d)^{-1}$, где g_d — плотность состояний в пространстве соответствующей размерности. Поэтому

$$c_i \approx \frac{J}{\Delta_i E} \approx \frac{J}{(g_d L^d)^{-1}} = J(g_d L^d). \quad (6.2)$$

Если интеграл перекрытия J мал, то коэффициенты c_i малы, и волновые функции остаются локализованными в исходных объемах L^d . Если же интеграл J велик, то волновые функции распространяются на весь объем $(2L)^d$. Отсюда следует предположение, что поведение волновых функций при удвоении размера куба L можно описать одним параметром (6.2).

Нетрудно видеть, что и транспортные свойства объема L^d меняются с величиной J точно так же. При большом J кондактанс велик, потому что электроны делокализованы и заряд может перейти с одной грани куба на противоположную путем металлической проводимости. При малом J проводимость мала; она реализуется за счет прыжков между центрами локализации или прямого туннелирования с одной грани гиперкуба на противоположную. Чем меньше J , тем меньше вероятность прыжков и туннелирования. Поэтому кондактанс можно считать той физически измеряемой величиной, которая позволяет судить о волновых функциях электронов в основном состоянии. Причем именно кондактанс, а не проводимость, потому что он безразмерен в единицах $e^2/\hbar$ при любой размерности пространства d , и потому что эта физическая величина сохраняет смысл сколь угодно близко к переходу.

Заметьте. Если оценить плотность состояний через ширину энергетической зоны W и электронную концентрацию, $g_d \simeq n_d/W$ (ср. с (5.1)), то параметр (6.2) по сути дела превратится в параметр Андерсона (5.3). Это еще раз указывает на связь между спецификой волновых функций и кондактансом.

В основополагающей работе [4], где была предложена однопараметрическая скейлинговая теория перехода металл–изолятор, в качестве функции состояния использован именно кондактанс образца y . Следуя работе [4], изобразим на рис. 6.1 логарифмическую производную кондактанса по размеру

$$\beta = \frac{d \ln y}{d \ln L} = \frac{L}{y} \frac{dy}{dL}$$

как функцию от аргумента $\ln y$:

$$\frac{d \ln y}{d \ln L} = f(\ln y). \quad (6.3)$$

Заметьте. Смысл скейлинговой функции β можно уяснить себе следующим образом. Если y — степенная функция L , например,

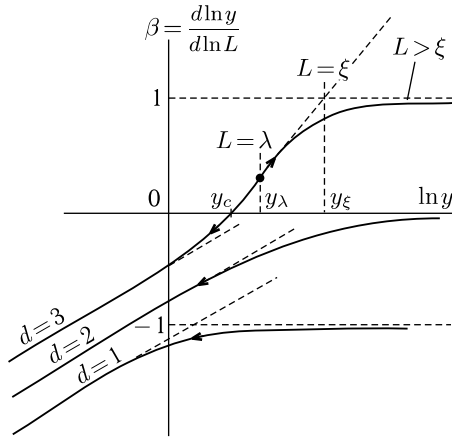


Рис. 6.1. Универсальные функции $\beta(\ln y)$ для различных размерностей [4]

$y = \kappa L^{d-2}$, то β — показатель степени: $\beta = d - 2$, а на графике на рис. 6.1 получится горизонтальная прямая. Именно такие степенные функции фигурируют в соотношениях (6.1).

Каждый конкретный образец размерности d изображается в виде точки на кривой $\beta_d(y)$. Согласно скейлинговой гипотезе функция $\beta(\ln y)$ универсальна для каждой размерности d , а состояния определяются только одним параметром y . По сути дела, кондактанс y выполняет роль управляющего параметра x , о котором говорилось выше.

Асимптотики универсальных функций $\beta_d(y)$ при очень больших и очень малых y можно получить из общих соображений. При больших y можно воспользоваться макроскопическим соотношением (6.1) транспортной теории, откуда

$$\beta \rightarrow d - 2 \quad \text{при} \quad y \rightarrow \infty. \quad (6.4)$$

В пределе малых y , когда все электроны локализованы, и их волновые функции затухают на расстояниях a_B , конечный кондактанс при $T = 0$ осуществляется за счет экспоненциальных хвостов волновых функций на противоположных гранях куба, и потому

$$y = y_0 e^{-L/a_B}, \quad \ln y = \ln y_0 - \frac{L}{a_B}, \quad \beta = \ln \frac{y}{y_0}. \quad (6.5)$$

Если соединить асимптотики (6.4) и (6.5) плавными линиями без экстремумов, то получаются кривые, представленные на рис. 6.1. Левые асимптотики этих кривых при $\ln y \rightarrow -\infty$ все лежат в нижней полуплоскости $\beta < 0$. Правые асимптотики при $\ln y \rightarrow +\infty$ расположены относительно оси абсцисс $\beta = 0$ по-разному. При $d < 2$ они лежат в нижней, а при $d > 2$ — в верхней полуплоскости. Эта разница

принципиальная, потому что при $d > 2$ кривая $\beta_d(\ln y)$ пересекает ось абсцисс, а наличие точки пересечения, как мы увидим ниже, означает существование в системе перехода металл–изолятор.

Отмеченная разница дает основания говорить, что двумерный случай является пограничным. Пограничные свойства двумерных систем стимулировали появление в теории переходов металл–изолятор так называемых « $(2 + \varepsilon)$ -моделей», где d считается непрерывной переменной, по которой производится разложение в окрестности значения $d = 2$.

Размер L является текущим параметром вдоль кривых $\beta_d(\ln y)$. Стрелки на кривых показывают, куда смещается изображающая точка при увеличении размера образца L : в верхней полуплоскости она смещается направо, а в нижней — налево. Корреляционная длина ξ , введенная формально в начале этой главы, определяет критический размер, отделяющий область, в которой справедливо классическое соотношение (6.1). При $d = 3$ это минимальный размер, при котором справедливо (6.1). Действительно, из кривой $d = 3$ на рис. 6.1 видно, что пока $L \ll \xi$, величина β существенно меняется при изменении L , а в области $L \gg \xi$ изображающая точка выходит на асимптотику, где $\beta = 1$ и от L уже не зависит. При $d = 1, 2$ это максимальный размер в том смысле, что при $L > \xi$ начинаются процессы локализации.

Кривые $\beta(y)$ предлагаются в качестве конечного продукта скейлинговой гипотезы. Воспользоваться этим продуктом можно следующим образом. Возьмем кубик (квадратную пленку, проволоку единичной длины) размера L из интересующего нас материала и мысленно измерим его кондактанс при абсолютном нуле температуры. Получим точку на одной из кривых на рис. 6.1. Если эта точка находится на нижней полуплоскости $\beta < 0$, то наш материал является изолятором: при увеличении L кондактанс y кубика становится экспоненциально малым. Если же точка находится на верхней полуплоскости $\beta > 0$, то мы имеем дело с металлом, кондактанс которого y растет с размером L .

При всей фантастичности такой процедуры скейлинговая гипотеза сыграла важную роль, позволив связать в единую стройную картину ряд разрозненных фактов. По поводу самой гипотезы можно задать несколько вопросов.

— Из чего состоит полный набор физических соображений, позволивший сформулировать эту гипотезу?

— Существуют ли теоретические вычисления, хоть в какой-то мере ее подтверждающие?

— Каковы физические следствия скейлинговой гипотезы и допускают ли они экспериментальную проверку?

То, что сказано выше о проводимости, кондактансе и асимптотиках, является лишь частью ответа на первый вопрос. Мы, однако, этим ограничимся, отослав интересующихся к оригинальной статье [4] и обзору [2]. Отметим лишь одно обстоятельство. *На всех этапах рассуждений, например при написании выражения (6.5), подразумевается волновая функция одного электрона и никак не учитываются ни со-*

стояние всех остальных электронов, ни спинные взаимодействия. Поэтому кривые на рис. 6.1, строго говоря, относятся только к системе невзаимодействующих бесспиновых электронов. Обычно заранее не ясно, в какой мере ту или иную электронную систему можно считать невзаимодействующей. Поэтому общая стратегия заключается в том, чтобы, сравнивая с предсказаниями гипотезы результаты измерений для разных электронных систем, постараться выяснить применимость гипотезы, ее универсальность и роль взаимодействий.

Перейдем ко второму вопросу. Возможности сравнения скейлинговой гипотезы с теоретическими вычислениями ограничены, но все же существуют. Переход от свободных носителей к полностью локализованному, описываемый кривыми $\beta_d(y)$ на рис. 6.1, начинается со слабой локализации. Поэтому спроектируем на эти кривые существующие теоретические расчеты квантовой поправки к проводимости при слабой локализации. При $T = 0$ неупругая длина $L_\varphi = \infty$. Полагая, что электрон теряет память о фазе, достигнув поверхности образца и рассеявшись на ней, подставим L во все формулы теории слабой локализации вместо L_φ . Получающиеся поправки к функциям $\beta_d(y)$ для всех трех размерностей сведены в табл. 6.1, где через σ_d ($d = 1, 2, 3$) обозначена больцмановская проводимость в d -мерном пространстве.

Таблица 6.1

	$d = 1$	$d = 2$	$d = 3$
σ	$\sigma_1 + K_1 l - K_1 L$	$\sigma_2 - K_2 \ln \frac{L}{l}$	$\sigma_3 - \frac{K_3}{l} + \frac{K_3}{L}$
y	$\frac{\hbar}{e^2} \left(\frac{\sigma_1}{L} - K_1 \right)$	$\frac{\hbar}{e^2} \left(\sigma_2 - K_2 \ln \frac{L}{l} \right)$	$\frac{\hbar}{e^2} (\sigma_3 L + K_3)$
β	$\frac{L}{y} \left(-\frac{\hbar}{e^2} \frac{\sigma_1}{L^2} \right) = \frac{-\sigma_1}{\sigma_1 - K_1 L}$	$\frac{1}{y} \frac{dy}{d \ln L}$	$\frac{\hbar}{e^2} \sigma_3 \frac{L}{y} = \frac{1}{1 + K_3/\sigma_3 L}$
β_d	$-1 - K_1/Y$	$0 - K_2/Y$	$1 - K_3/Y$
Константы K_1 , K_2 и K_3 имеют размерность $[\text{Ом}^{-1}]$			

Вычисленные в предпоследней строке таблицы функции β представлены в последней строке в унифицированном виде $\beta_d = (d - 2) - K_d/Y$. Они имеют правильную асимптотику при $Y \rightarrow \infty$ и отклоняются от нее вниз при конечных Y , как это предполагается гипотезой и как это нарисовано на рис. 6.1. Таким образом там, где оказалось возможно провести расчет в рамках теории возмущений, его результат подтвердил скейлинговую гипотезу.

Обратите внимание. Последнее утверждение не столь бесспорно, как кажется. Оно исходит из знака квантовых поправок к проводимости (2.11) и по существу означает следующее: слабая локализация является первым шагом на пути к сильной, настоящей локализации. Но ведь интерференционная квантовая поправка

устроена довольно хитро, и при наличии достаточно сильного спин-орбитального взаимодействия слабая локализация уступает место антилокализации. Как совместить скейлинговую гипотезу и спин-орбитальное взаимодействие? Мы вернемся к обсуждению этого в конце главы, после того как станет ясно, где это может быть наиболее существенно.

А сейчас перейдем к третьему вопросу — к конкретным следствиям скейлинговой гипотезы.

6.2. Трехмерные (3D) системы

Возьмем маленький пробный кубик из исследуемого материала размером λ . Если точка, отображающая на скейлинговую кривую β_3 -кондактанс y_λ этого кубика, лежит в нижней полуплоскости $\beta < 0$, то увеличение размера куба $L > \lambda$ сдвигает отображающую точку влево. Увеличивая L , можно сделать кондактанс сколь угодно малым. (*Заметьте:* при $T = 0$!) Этот материал — изолятор. Если же отображающая точка $\beta(y_\lambda) > 0$, то увеличение размера L сдвигает отображающую точку вправо, так что кондактанс можно сделать сколь угодно большим. Это — металл. Пересечение кривой $\beta_3(y)$ в точке y_c осью $\beta = 0$ означает существование перехода металл–изолятор.

Аппроксимируем 3D скейлинговую кривую $\beta_3(y)$ в верхней полуплоскости $\beta > 0$ ломаной, состоящей из отрезка, выходящего из точки $(\ln y_c, 0)$ под наклоном $s = (d\beta/d \ln y)_{y_c}$ до пересечения с прямой $\beta = 1$ (обычно считают, что s порядка единицы), и горизонтальной полупрямой $\beta = 1$. Кондактанс в точке излома обозначим через y_ξ , а соответствующий размер L через ξ (рис. 6.1). Уравнение отрезка

$$\frac{d \ln y}{d \ln L} = s \ln \frac{y}{y_c} \quad (6.6)$$

в силу специфичности функции β является дифференциальным уравнением, общее решение которого имеет вид

$$\ln \frac{y}{y_c} = \left(\frac{L}{\lambda} \right)^s \ln \frac{y_\lambda}{y_c}. \quad (6.7)$$

Здесь λ играет роль начального условия, заданного в одной точке отрезка. При $y = y_\xi$ в точке излома,

$$\ln \frac{y_\xi}{y_c} = \frac{1}{s}. \quad (6.8)$$

Размер куба, который приводит изображающую точку в точку излома, обозначим через ξ . Из (6.7) следует, что

$$\xi = \lambda \left(s \ln \frac{y_\lambda}{y_c} \right)^{-1/s}. \quad (6.9)$$

При дальнейшем увеличении размера $L > \xi$ отображающая точка будет двигаться вдоль горизонтальной части кривой, где существует прово-

димность σ и можно пользоваться выражением (6.1) для y . Так как ξ — наименьшая длина, при которой имеет смысл понятие удельной проводимости, она называется корреляционной длиной. Наконец, подставив значение (6.8) для y_ξ и значение $L = \xi$ в выражение (6.1), получим окончательно

$$\sigma = \left(\frac{e^2}{\hbar}\right) \frac{y_\xi}{\xi} = \left(\frac{e^2}{\hbar}\right) \frac{\ln y_c + 1/s}{\xi} \sim \left(\frac{e^2}{\hbar}\right) \frac{1}{\xi}. \quad (6.10)$$

Размер λ пробного куба не может быть сколь угодно малым. Минимальный разумный размер — это электронная длина волны $\lambda \approx 1/k_F$. Если кондактанс пробного куба y_λ попадет в интервал $y_c < y_\lambda < y_\xi$, то корреляционная длина окажется $\xi > l_{\min} \approx k_F^{-1}$ и $\sigma < \sigma_{\text{Mott}}$.

Если y_λ близко к y_c , то $\ln y_\lambda/y_c \approx (y_\lambda/y_c) - 1$. Положив для оценки $s = 1$, получим из (6.9)

$$\xi \approx \lambda \frac{y_c}{y_\lambda - y_c}. \quad (6.11)$$

При $y_\lambda \rightarrow y_c$ величина $\xi \rightarrow \infty$, а проводимость $\sigma \rightarrow 0$. Таким образом из скейлинговой гипотезы следует, что 3D удельная проводимость достаточно большого образца может быть сколь угодно малой, так что переход металл–изолятор является непрерывным.

Заметьте. Маленький пробный кубик нам понадобился только на промежуточной стадии рассуждений. Заключения относительно проводимости σ и характера перехода относятся к макроскопическим образцам.

Ценность кривой $\beta(y)$ в степени ее универсальности, которая в первую очередь существенна в окрестности перехода. Здесь она выражается в том, что кривая $\beta(y)$ должна иметь для разных систем один и тот же наклон s . Значение s характеризует так называемый класс универсальности. Для невзаимодействующих электронов из $(2 + \varepsilon)$ -модели следует, что наклон $s = \varepsilon$. Если предположить, что разложением $d = 2 + \varepsilon$ можно пользоваться вплоть до $\varepsilon = 1$, то для трехмерных систем получается $s = 1$.

Как мы знаем из гл. 2, при конечной температуре $T \neq 0$ размер области фазовой когерентности задается длиной L_φ , которая определяется процессами электрон-фононного или электрон-электронного рассеяния, т. е. это минимальная из длин (2.7) и (2.30). Это позволяет использовать кривые фазовой диаграммы рис. 6.1 и при конечных температурах, заменяя L на L_φ . Например, если система изначально имеет размер $L = \infty$, но находится при конечной температуре, то область когерентности, которая могла бы установиться благодаря близости к переходу, ограничивается не размером системы, а температурой. Поэтому переменная L в уравнении (6.3) заменяется на длину сбоя фазы L_φ . При уменьшении температуры изображающая точка движется вдоль соответствующей кривой $\beta_d(y)$ так же, как при увеличении L . Выше мы уже использовали это при сравнении фазовых диаграмм на рис. 6.1 с расчетами квантовых поправок.

В интервале от x_c до x_m проводимость, согласно (6.10), выражается через корреляционную длину ξ , а при $x > x_m$ работает формула Друде. Суммируя, можно написать, что при $T = 0$

$$\sigma(0) = \left(\frac{e^2}{\hbar}\right) \cdot \begin{cases} 0, & x \leq x_c, \\ 1/\xi, & x_c < x \leq x_m, \\ (k_{\text{Fl}})^2/l, & x \geq x_m. \end{cases} \quad (6.13)$$

Сшивая два последних выражения при $x = x_m$, получаем, что в этой точке $\xi = l$.

По другую сторону от перехода, со стороны изолятора, выделенной является точка x_i , в которой монотонно падающая по мере удаления от перехода корреляционная длина ξ сравнивается с боровским радиусом a_B . По сути дела корреляционная длина имеет физический смысл только в интервале $x_i < x < x_m$, где ξ остается больше длины свободного пробега l (справа) или боровского радиуса a_B (слева) и замещает эти величины в формулах для проводимости.

При описании проводимости при $T = 0$ мы воспользовались скейлинговой гипотезой для невзаимодействующих электронов и формулой (6.13), но при анализе окрестности перехода при конечных температурах будем исходить из квантовых поправок, обусловленных межэлектронным взаимодействием. Это дает основание называть переход в x_c переходом Андерсона-Мотта. Обратимся к фазовой плоскости (x, T) на рис. 6.2 и начнем с самой правой части металлической области, где справедливо больцмановское выражение σ_3 для классической трехмерной проводимости с квантовой поправкой:

$$\sigma = \sigma_3 + \frac{e^2}{\hbar} L_{ee}^{-1}, \quad L_{ee} = \sqrt{\hbar D/T}. \quad (6.14)$$

После пересечения линии $x = x_m$, которую называют пределом Мотта, формула (6.14) сменяется на

$$\sigma = \frac{e^2}{\hbar} \left(\frac{1}{\xi} + \frac{1}{L_{ee}} \right) = \frac{e^2}{\hbar} \left(\frac{1}{\xi} + (T/\hbar D)^{1/2} \right). \quad (6.15)$$

При достаточно низких температурах, пока квантовая поправка относительно мала: $1/\xi \gg 1/L_{ee}$, первый член суммы в формуле (6.15) зависит только от беспорядка x , а второй только от температуры. Однако, по мере приближения к переходу (или при повышении температуры) это неравенство нарушается, и мы попадаем в область, называемую критической. В ней теория квантовых поправок уже не работает, и формулу (6.15) следует считать интерполяционной. Чтобы учесть особенность критической области, предположим, что коэффициент диффузии D зависит в ней не только от x , но и от T : диффузия происходит как бы на тех самых флуктуациях электромагнитного поля, которые определяют L_{ee} , и потому становится температурозависящей.

Тогда можно написать систему уравнений для функций $\sigma(T)$ и $D(T)$, используя в качестве второго уравнения соотношение Эйнштейна:

$$\begin{cases} \sigma = \frac{e^2}{\hbar} \left(\frac{1}{\xi} + \sqrt{\frac{T}{\hbar D}} \right), \\ \sigma = e^2 g_F D \end{cases} \quad (6.16)$$

(g_F – плотность состояний на ферми-уровне).

Исключим D из этих двух уравнений. Затем разрешим оставшееся уравнение относительно $\sigma(T)$, воспользовавшись тем, что мы находимся достаточно близко к переходу и что $1/\xi \ll 1/L_{ee}$. Получим температурную зависимость проводимости в правой части критической области $x > x_c$

$$\sigma(T) = \frac{e^2}{\hbar} \left(\frac{1}{\xi} + (T g_F)^{1/3} \right) \equiv \alpha(x) + \beta T^{1/3}. \quad (6.17)$$

Точно на переходе $\xi = \infty$ и $\alpha = 0$.

Граница критической области, которая определяется из сравнения двух слагаемых в правой части уравнения (6.17), описывается кубической параболой в осях $(1/\xi, T)$:

$$T \propto (g_F \xi^3)^{-1}. \quad (6.18)$$

Итак, металлическая область состоит из трех подобластей: вдали от перехода проводимость определяется формулой Друде с квантовыми поправками (6.14), вблизи перехода при низких температурах она контролируется корреляционной длиной ξ с относительно малой квантовой поправкой — формула (6.15), а при несколько более высоких температурах, в критической области, она описывается формулой (6.17).

Заметьте. Степень $1/3$ возникла благодаря тому, что в качестве времени сбоя фазы τ_φ было взято время расфазировки $\tau_{ee} \propto T^{-1}$.

При ином механизме сбоя фазы с $\tau_\varphi \propto T^{-\nu}$, температура войдет в уравнение (6.17) в степени $\nu/3$.

В области изолятора тоже есть три подобласти. Вдали от перехода проводимость определяется одним из прыжковых механизмов, описанных в гл. 4; согласно формулам (4.18), (4.19) (либо формулам (4.28), (4.29)), прыжковая проводимость определяется отношением средней длины прыжка $\bar{r}$ к длине затухания локализованной волновой функции a_B :

$$\sigma \propto \exp(-\bar{r}/a_B), \quad x < x_i. \quad (6.19)$$

При приближении к переходу, правее точки x_i , где $\xi > a_B$, длина затухания увеличивается, вместо a_B она становится равной ξ , и формула (6.19) для проводимости изменяется:

$$\sigma \propto \exp(-\bar{r}/\xi), \quad x_i < x < x_c. \quad (6.20)$$

При этом следует иметь в виду, что в области $\xi > a_B$ средняя длина прыжка $\bar{r}$ тоже растет по мере приближения к переходу, но медленнее, чем ξ .

Наконец, есть еще и критическая область, которая должна существовать не только справа, но и слева от точки перехода. Естественно думать, что при пересечении линии $x = x_c$ степень $T^{1/3}$ в температурной зависимости не должна измениться. Однако слагаемое $\alpha(x)$ в выражении (6.17), как мы покажем, станет отрицательной. Положение левой границы критической области определим из сравнения корреляционной длины ξ со средней длиной прыжка, $\xi = \bar{r}$. Используя выражение (4.18) для $\bar{r}$, в которое в качестве a_B подставлена корреляционная длина ξ , получим уравнение кубической параболы (6.18). На левой границе критической области критическая проводимость σ_{crit} должна перейти в прыжковую проводимость σ_{hopp} , существенно меньшую по величине. Пренебрегая второй по сравнению с первой, можно считать, что на этой границе $\sigma_{\text{crit}} \approx 0$ и

$$\alpha \approx -\beta T^{1/3} \approx \frac{e^2}{\hbar} \frac{1}{\xi}. \quad (6.21)$$

Объединяя уравнения (6.17) и (6.21), получим, что проводимость в критической области

$$\sigma_{\text{crit}} = \alpha + \beta T^{1/3}, \quad \alpha = \frac{e^2}{\hbar} \cdot \begin{cases} 1/\xi, & x > x_c, \\ -1/\xi, & x < x_c, \end{cases} \quad \beta = \frac{e^2}{\hbar} g_F^{1/3}. \quad (6.22)$$

Заметьте. Из металлического состояния в состояние изолятора (оба при $T = 0$) можно перейти, обогнув точку фазового перехода, см. рис. 6.3. При этом, однако, придется пересечь две линии кроссоверов, в окрестности которых сменяются доминирующие процессы, определяющие проводимость. На последнем этапе такой траектории проводимость будет убывать экспоненциально с понижением температуры.

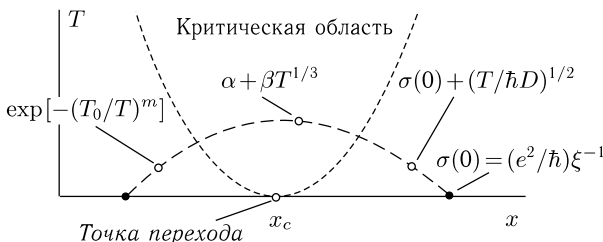


Рис. 6.3. Обход точки фазового перехода через критическую область

Для экспериментального наблюдения перехода металл–изолятор нужно выбрать управляющий параметр, измерить температурные зависимости проводимости при разных значениях этого параметра, постро-

ить их в спрямляющих осях ($T^{1/3}, \sigma$) и, убедившись, что получаются прямые линии, проэкстраполировать эти прямые к $T = 0$ для определения $\sigma(x, 0)$. Переход должен проявиться в поведении функции $\sigma(x, 0)$ в точке $x = x_c$.

Самым естественным управляющим параметром является концентрация носителей n , обусловленная количеством примесей в данном образце. Каждая экспериментальная кривая на рис. 6.4 соответствует измерениям вдоль вертикальной прямой на диаграммах рис. 6.2 или рис. 6.3. Судя по экстраполяции экспериментальных данных к $T = 0$, для Ge:As критическая концентрация As равна $n_c = 3,5 \cdot 10^{17} \text{ см}^{-3}$: при концентрациях $n > n_c$ экстраполяция дает положительные значения $\sigma(0)$. На самой нижней зависимости $\sigma(T)$, отвечающей концентрации $n = 3,0 \cdot 10^{17} \text{ см}^{-3}$, стрелкой отмечена температура, вблизи которой происходит кроссовер из критической области в область прыжковой проводимости. Прыжковая проводимость существенно меньше, чем проводимость в критической области, и ее практически не видно при использованном на рис. 6.4 масштабе, так что зависимость $\sigma(T)$ как бы упирается в ось абсцисс. Это было использовано выше при написании граничного условия (6.21).

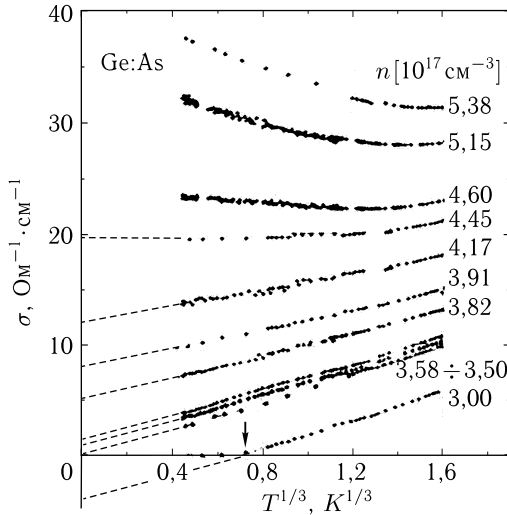


Рис. 6.4. Температурная зависимость проводимости серии образцов Ge:As с разным уровнем легирования в районе перехода металл-изолятор, [5]

Как видно из рис. 6.4, у серии кривых $\sigma(T)$ на Ge:As есть еще одно выделенное значение концентрации примесей: при $n = 4,45 \cdot 10^{17} \text{ см}^{-3}$ меняет знак производная $d\sigma/dT$ при низких температурах. Естественно считать, что это и есть то значение управляющего параметра $x_m = n_m$, при котором выполняется соотношение (6.12), потому что при этой

концентрации имеющие разные знаки классическая и квантовая температурозависящие поправки к проводимости компенсируют друг друга. Это означает, что приблизительно при этой концентрации друдевское описание проводимости перестает работать, и начинается критическая область (см. также соотношение (6.13)). Знание n_m означает знание $\sigma_{\text{Mott}} \equiv \sigma(n_m) \approx 20 \Omega^{-1} \text{ см}^{-1}$, и абстрактная диаграмма на рис. 6.2 обрывается конкретными числами: выясняется отношение n_m/n_c , определяется коэффициент пропорциональности между $1/\sigma$ и ξ в интервале концентраций $n_c < n < n_m$, а проводимость можно записать в приведенных единицах $\sigma/\sigma_{\text{Mott}}$.

Образец Ge:As с концентрацией $n = 4,6 \cdot 10^{17} \text{ см}^{-3}$ находится, согласно рис. 6.4, в глубине металлической области. Но его можно сделать изолятором при помощи магнитного поля, которое в данном случае выполняет роль управляющего параметра. Как видно из рис. 6.5, переход в состояние изолятора произошел в поле 5 Тл. В более сильных

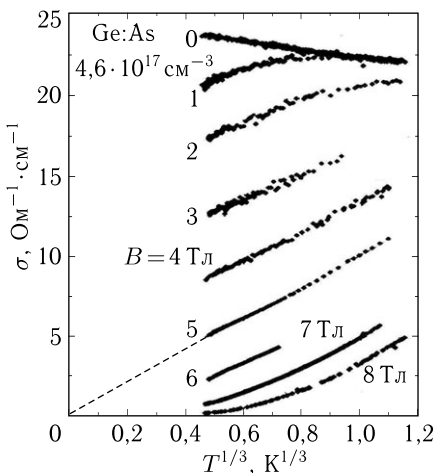


Рис. 6.5. Температурная зависимость проводимости одного образца Ge:As в разных магнитных полях [6]

полях, 7 и 8 Тл, экспериментальные точки при самых низких температурах отклоняются от прямых. Это результат приближения к области прыжковой проводимости (см. рис. 6.2) — проводимость σ не может стать отрицательной (на рис. 6.4 аналогичный изгиб в нулевом поле при $n = 3 \cdot 10^{17} \text{ см}^{-3}$ выглядит как излом).

Зачастую точности измерений проводимости и диапазона изменений температуры недостаточно, чтобы однозначно определить функциональную зависимость $T^{1/3}$. Иногда при обработке экспериментальных данных используют для спрямления кривых не $T^{1/3}$, а $T^{1/2}$, используя вплоть до перехода степень, определяющую квантовую поправку

в больцмановской области при $\Delta\sigma \ll \sigma(0)$. На определении значения управляющего параметра в точке перехода это почти не сказывается.

Таким образом, эволюция транспортных свойств в окрестности перехода металл–изолятор вполне соответствует описанию, базирующемуся на скейлинговой гипотезе. Однако, формулы для проводимости (6.15) и (6.17), а также скейлинговое описание ничего не говорят о волновых функциях и механизмах проводимости в области (x_c, x_m) (см. рис. 6.2). Мы обсудим этот вопрос в следующем параграфе.

6.4. Мультифрактальные волновые функции ¹⁾

В интервале $[x_i, x_m]$ значений управляющего параметра x на фазовой плоскости (x, T) на рис. 6.2, т. е. в окрестности перехода металл–изолятор, в формулы для проводимости входит корреляционная длина ξ . Это феноменологическая величина, введенная из скейлинговых соображений. То, что в этом интервале требуют модернизации и формула проводимости (1.5) кинетической теории, и формулы (4.18), (4.28) прыжковой проводимости, означает, что электроны в этом интервале приобретают специфические свойства. Эти свойства сейчас описываются при помощи теории мультифрактальности.

Рассмотрим волновую функцию $\psi(\mathbf{r})$ электрона, делокализованного в ограниченном трехмерном пространстве (металл, размерность $d=3$, размер L^3). Разобьем этот объем на мелкие ячейки объемом s^d . Пусть сначала $\psi(\mathbf{r})$ — гладкая функция. Благодаря нормировке $\int_{L^d} |\psi(\mathbf{r})|^2 d\mathbf{r} = 1$, вероятность P того, что электрон находится в малом объеме s^d , зависит от L :

$$P_L = |\psi(\mathbf{r})|^2 s^d \propto L^{-d}. \quad (6.23)$$

Постулируется, что если система находится достаточно близко к переходу металл–изолятор, то из-за случайного потенциала плотность вероятности $|\psi(\mathbf{r})|^2$ становится функцией, сильно флуктуирующей на расстояниях, много меньших корреляционной длины ξ . Тогда при росте L , но при условии $L < \xi$, мы можем обнаружить, что для конкретной i -й ячейки

$$P_L = |\psi(\mathbf{r}_i)|^2 s^d \propto L^{-\alpha(\mathbf{r}_i)}, \quad \alpha \neq d, \quad (6.24)$$

причем α меняется от ячейки к ячейке. Это означает, что рост L в интервале $L < \xi$ приводит к неоднородному уменьшению $|\psi(\mathbf{r})|^2$ и к нарастанию относительной величины флуктуаций. Именно такое поведение (если оно действительно существует, в чем нам еще пред-

¹⁾ С математической основой теории мультифрактальных функций проще всего познакомиться в книге [7]; применения этих математических понятий к теории переходов металл–изолятор подробно описаны, например, в обзорах [8, 9].

стоит убедиться) и описывается теорией мультифрактальных функций. Простейший математический пример построения мультифрактальной функции и анализ ее свойств имеются в книге [7] (гл. 6).

Возможные значения α образуют целый интервал от $\alpha_{\min} < d$ до $\alpha_{\max} > d$, и для любого α из этого интервала при данном L имеется объем V_α , состоящий из большого количества отдельных частей, в котором реализуется соотношение (6.24) с данным α . Объем V_α изменяется как некоторая степень L :

$$V_\alpha \propto L^{f(\alpha)}, \quad \int_{\alpha_{\min}}^{\alpha_{\max}} V_\alpha d\alpha = L^d. \quad (6.25)$$

Это означает, что d -пространство разбивается на области V_α с фрактальными размерностями $f(\alpha)$. Функция $f(\alpha)$ обращается в ноль на краях интервала $[\alpha_{\min}, \alpha_{\max}]$ и достигает максимума $f(\alpha_0) = d$ в точке $\alpha_0 > d$.

При анализе свойств собственных функций гамильтониана со случайным потенциалом полезны интегралы степеней q квадрата волновой функции. Для гладких $|\psi(\mathbf{r})|^2$

$$I_q = \int_L |\psi(\mathbf{r})|^{2q} d\mathbf{r} \propto L^{d-qd} = L^{-(q-1)d}. \quad (6.26)$$

Для мультифрактальных функций, имеющих сингулярности в каждой точке, в качестве подынтегральной функции в интеграле (6.26) можно взять среднее от величины P_L по всем ячейкам s^d в объеме L^d :

$$I_q = \int_L \langle |\psi(\mathbf{r})|^{2q} \rangle d\mathbf{r} \propto L^{-\tau(q)} = L^{-(q-1)D_q}. \quad (6.27)$$

Запись скейлингового показателя $\tau(q)$ в виде $(q-1)D_q$ подчеркивает, что соотношение (6.27) является обобщением соотношения (6.26). Функция $D_q(q)$ называется спектром обобщенных фрактальных размерностей функции $|\psi(\mathbf{r})|^2$.

Интеграл I_q мультифрактальной функции можно оценить и иным способом. Объем областей $V_\alpha \propto L^{f(\alpha)}$, а функция $|\psi(\mathbf{r})|^2$ в этих областях, согласно соотношению (6.24), пропорциональна $L^{-\alpha}$. Поэтому

$$I_q \propto L^{-\tau(q)} \propto \int_{\alpha_{\min}}^{\alpha_{\max}} L^{-q\alpha + f(\alpha)} d\alpha. \quad (6.28)$$

С помощью этого уравнения можно установить связь между $\tau(q)$ и $f(\alpha)$:

$$f(\alpha) = \alpha q - \tau(q), \quad \alpha = \frac{d\tau(q)}{dq}. \quad (6.29)$$

То, что соотношения (6.24), (6.27), (6.28) действительно реализуются вблизи перехода Андерсона и что $D_q \neq d$, было многократно продемонстрировано в аналитических и численных исследованиях, исходящих непосредственно из гамильтониана со случайным потенциалом. Из них следует, что даже с металлической стороны вблизи перехода металл–изолятор волновые функции вблизи уровня Ферми имеют многочисленные сингулярности, так что области с провалами и области с пиками функции $|\psi(\mathbf{r})|^2$ соседствуют на малых расстояниях. При этом сохраняются корреляции между значениями $\psi(\mathbf{r})$ на расстояниях, меньших ξ ($\xi > k_F^{-1}$), даже у функций $\psi(\varepsilon)$ с близкими энергиями ε , потому что эти значения определяются в конечном счете одним и тем же случайным потенциалом. Это сказывается на матричных элементах, определяющих, например, вероятности переходов между состояниями.

Современная экспериментальная техника позволяет непосредственно просканировать значения функции $|\psi(\mathbf{r})|^2$ вдоль некоторой плоскости, перемещая вдоль нее иглу сканирующего электронного микроскопа и измеряя локальную плотность электронных состояний $g_l(\varepsilon)$, пропорциональную дифференциальной проводимости $dI/dV(\varepsilon)$ при заданном постоянном потенциале $V = \varepsilon/e$ на игле (здесь энергия ε отсчитывается от уровня Ферми; см. Приложение Б). На рис. 6.6 при-

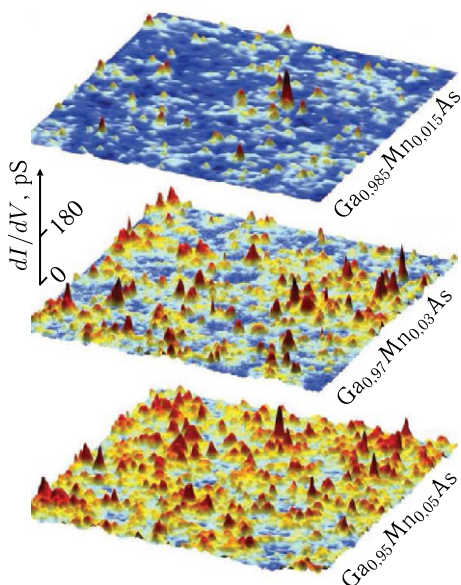


Рис. 6.6. Дифференциальная проводимость dI/dV на ферми-уровне вдоль поверхности эпитаксиально выращенных образцов $\text{Ga}_{1-x}\text{Mn}_x\text{As}$ с разными концентрациями x , измеренная иглой сканирующего туннельного микроскопа.

Площадь сканирования 700 Å на 700 Å [10]

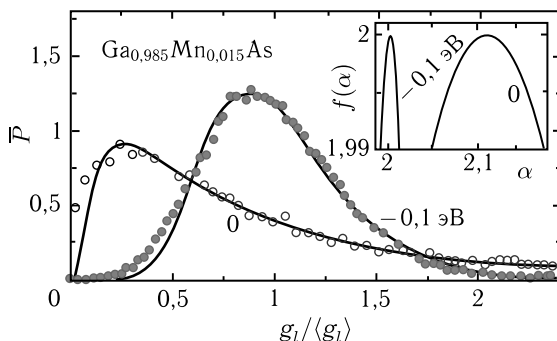


Рис. 6.7. Распределение $\bar{P}$ для образца с концентрацией Мп 1,5% (пустые кружочки); сплошная кривая — log-нормальное распределение, около кривой указана отсчитанная от ε_F энергия ε , при которой проведено измерение, $\varepsilon = 0$. Для сравнения приведено аналогичное распределение для того же образца, но при $\varepsilon = -0,1$ эВ (заполненные кружочки); сплошная кривая — нормальное распределение. На вставке: рассчитанные по экспериментальным данным мультифрактальные спектры $f(\alpha)$ вблизи точек максимума α_0 [10]

ведены данные, полученные при сканировании поверхности образцов $\text{Ga}_{1-x}\text{Mn}_x\text{As}$ толщиной 2000 Å, выращенных методом молекулярной эпитаксии на поверхности GaAs, с концентрацией Мп, близкой к критической для перехода металл–изолятор [10]. Эти данные о локальной плотности состояний g_l на ферми-уровне не только наглядно иллюстрируют крайнюю степень ее неоднородности, но и позволяют получить количественные характеристики этой неоднородности. Для примера на рис. 6.7 приведены результаты обработки данных для образца с концентрацией магния $x = 1,5\%$. Распределение вероятности $\bar{P}$ конкретных значений величины дифференциальной проводимости dI/dV на ферми-уровне, нормализованное на ее среднее значение на просканированной площади, описывается не нормальным распределением, симметричным относительно среднего значения $\langle g_l \rangle$, а log-нормальным, с максимумом, сдвинутым в область малых g_l , и хвостом в области больших g_l . Такое обычно получается, когда обладающая дисперсией величина находится в показателе степени выражения, функция распределения которого изучается.

Для сравнения на рис. 6.7 представлена кривая $\bar{P}(g_l/\langle g_l \rangle)$, полученная из аналогичного массива данных на том же образце, $x = 1,5\%$, но в глубине валентной зоны, т.е. при $\varepsilon = -0,1$ эВ. Она описывается нормальным распределением. Сравнение данных трех массивов, представленных на рис. 6.6, показывает, что по мере увеличения x и удаления от перехода металл–изолятор максимум на кривой $\bar{P}(g_l/\langle g_l \rangle)$ постепенно сдвигается вправо, приближаясь к единице: распределение из log-нормального постепенно становится нормальным.

Получаемые в эксперименте массивы данных позволяют вычислять и интегралы (6.27) и, следовательно, функции $\tau(q)$ и $f(\alpha)$. На вставке к рис. 6.7 приведены функции $f(\alpha)$ для обоих распределений с основной части рисунка. Хотя функции $|\psi(\mathbf{r})|^2$ заданы в трехмерном пространстве, мы имеем дело с двумерными массивами данных. В соответствии с этим, максимум функции $f(\alpha)$ в точке α_0 равен двум, сама точка α_0 совпадает с $d = 2$ для нормального распределения и сдвигается в область $\alpha_0 > 2$ по мере увеличения мультифрактальности.

Естественный вопрос, как влияет мультифрактальность волновых функций на транспорт вблизи перехода металл–изолятор, пока не имеет исчерпывающего ответа. Сравнительно простую качественную картину можно предложить на стороне изолятора. Из-за многочисленных узких «провалов» в локализованной волновой функции она становится рыхлой и расплывается, затухая на расстояниях $\xi > a_B$. Замена a_B на $\xi > a_B$ означает рост прыжковой проводимости, и причина этого роста именно в мультифрактальности. При движении изображающей точки в сторону перехода на рис. 6.2 проводимость остается прыжковой до тех пор, пока показатель экспоненты $\tilde{r}/\xi$ остается меньше единицы. Затем система переходит в критическую область.

Концепция мультифрактальности волновых функций позволяет придать физический смысл скейлинговой формуле (6.13) в непосредственной окрестности перехода на его металлической стороне: $\sigma = (e^2/\hbar)(1/\xi)$. Для этого заметим, что в точке x_m , на левом краю друдевской области, в которой проводимость следует обычной газовой модели, проводимость при $T = 0$ можно записать в виде $\sigma(x_m) = (e^2/\hbar)k_F$. Величина k_F не только характеризует количество $n \sim k_F^3$, участвующих в переносе электронов, но в то же время определяет, в силу свойств преобразования Фурье, размер волнового пакета u отдельного электрона в координационном пространстве,

$$u \sim k_F^{-1}. \quad (6.30)$$

В мультифрактальной области из-за рыхлости волновой функции размер пакета увеличивается и становится равным $u \sim \xi > k_F^{-1}$. Фурье-спектр такого пакета содержит компоненты с волновыми векторами k в интервале

$$0 < k < (k_F)_{\text{eff}}, \quad (k_F)_{\text{eff}} \equiv u^{-1} < k_F. \quad (6.31)$$

В спектре мультифрактальной функции имеются также компоненты с очень большими $k > k_F$, обеспечивающие ее локальную изрезанность, но они для нас здесь не представляют интереса, поскольку количество электронных волн $(k_F)_{\text{eff}}^3$, которые задают огибающую пакета, определяется именно интервалом $[0, (k_F)_{\text{eff}}]$. Применив соотношения (6.30) и (6.31), получим, что по мере приближения к переходу $\xi \rightarrow \infty$, а величина $n_{\text{eff}} \sim (k_F)_{\text{eff}}^3$ стремится к нулю. Другими словами, количество электронов, участвующих в переносе в мультифрактальной области, при приближении к переходу постепенно уменьшается до нуля.

6.5. Двумерные (2D) системы

Кривая $d = 2$ на рис. 6.1 вся находится в области отрицательных β . Из какого бы материала мы ни изготовили 2D систему, постепенное увеличение ее размеров должно приводить к уменьшению контактанса, так что контактанс станет экспоненциально малым и в пределе $L \rightarrow \infty$, $T \rightarrow 0$ должен обратиться в нуль. Другими словами, при достаточно низкой температуре любая пленка является изолятором. На первый взгляд это следствие скейлинговой гипотезы противоречит несомненному факту существования металлических пленок. Противоречие, однако, не является фатальным. Поскольку концентрация $n_2 \propto k_F^2$, а больцмановская 2D-проводимость $\sigma_2 = n_2 e^2 l / \hbar k_F \approx (e^2 / \hbar)(k_F l)$, проводимость 2D системы с учетом одной из квантовых поправок (2.11) или (2.37) имеет вид

$$\sigma = \sigma_2 - \frac{e^2}{\hbar} \ln \frac{L_\varphi}{l} \approx \frac{e^2}{\hbar} \left(k_F l - \ln \frac{L_\varphi}{l} \right). \quad (6.32)$$

Это выражение справедливо лишь до тех пор, пока поправка относительно мала. Но тенденция уменьшения проводимости с понижением температуры должна сохраниться. Поэтому формулу (6.32) можно использовать для оценки момента существенных изменений в характере проводимости, приравняв проводимость (6.32) нулю. Получившееся в результате значение L_φ назовем длиной локализации и обозначим буквой ξ :

$$\xi \approx l \exp(k_F l). \quad (6.33)$$

Важно, что так определенная длина ξ зависит только от электронной концентрации (через k_F) и от беспорядка (через l). Предположим, что логарифмическая квантовая поправка в (6.32) обусловлена межэлектронной интерференцией, так что частота расфазировки равна $\tau_{ee}^{-1} = T/\hbar$, диффузионная длина равна $L_{ee} = (\hbar D/T)^{1/2}$ и правая часть уравнения (6.32) обращается в ноль при температуре

$$T_{cr} = \frac{\hbar D}{\xi^2} \approx \frac{\hbar}{\tau} \exp(-2k_F l). \quad (6.34)$$

Заметьте. Если частота τ_{ee}^{-1} (или τ_φ^{-1}) меняется как температура в степени $\nu \neq 1$, т. е. $\tau_\varphi^{-1} \propto T^\nu$, то в показателе экспоненты в формуле (6.34) вместо 2 окажется $2/\nu$. Это существенно не изменит дальнейшие рассуждения.

Если бы при этой температуре формула (6.32) еще работала, то проводимость обратилась бы в ноль. Поэтому мы вправе ожидать, что в районе температуры (6.34) произойдет кроссовер от логарифмического падения проводимости (6.32) к экспоненциальному падению. Но даже при сравнительно небольшом «металлическом» значении параметра $k_F l > 5 \div 10$ эта температура становится нереально низкой, так что фактически локализации не происходит. К тому же кроссовер может не

наступить и из-за конечности размеров образца L : даже охладившись до температуры $T_{\text{сг}}$, мы не увидим кроссовер, если $L < \xi$. Линия кроссоверов изображена на рис. 6.8, где величина $1/\xi$ отложена вдоль оси абсцисс в качестве количественной меры беспорядка; это сделано на основании (6.33).

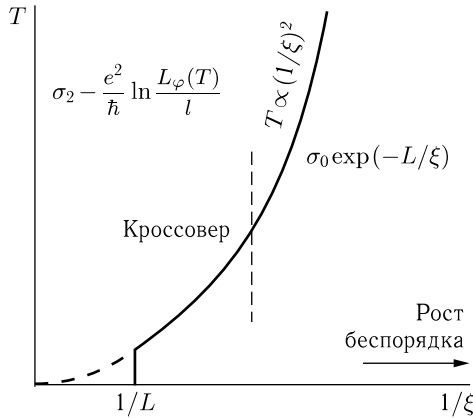


Рис. 6.8. Кроссовер от слабой к сильной локализации в двумерных системах при понижении температуры или росте беспорядка

Таким образом длина ξ , определяемая формулой (6.33), это действительно тот самый радиус локализации, на котором данный беспорядок в двумерной системе локализует носители при $T = 0$. В «металлических» пленках, $l \gg k_F^{-1}$, длина ξ экспоненциально велика, и локализация существует лишь гипотетически: для ее наблюдения нужны нереально большие размеры образцов $L \gg \xi$ и нереально низкие температуры (6.34). По мере роста беспорядка и уменьшения параметра $k_F l$ минимальный размер L падает, а температура (6.34) растет. При $k_F l \gtrsim 1$ оба параметра становятся реальными. Именно для этой области скейлинговая гипотеза вместо перехода металл-изолятор по мере увеличения беспорядка при $T = 0$ предсказывает для 2D-системы с фиксированным беспорядком кроссовер от слабой к сильной локализации при уменьшении температуры.

Представленные ниже эксперименты иллюстрируют изложенные соображения на различных 2D-системах. Такими системами могут быть либо особым образом изготовленные плоские поверхности раздела двух сред, вдоль которых образуется двумерная яма для электронов (гетероструктуры, инверсионные слои), либо тонкие пленки.

В гетероструктуре GaAs/Al_xGa_{1-x}As концентрацию 2D-электронов можно менять при помощи затвора: на одном образце при изменении электрического поля получается весь спектр концентраций. Как видно из рис. 6.9, функция $\rho(T)$ носит явно активационный характер при кон-

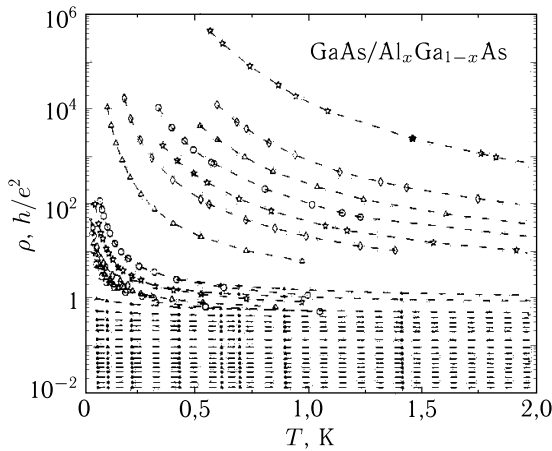


Рис. 6.9. Эволюция температурной зависимости сопротивления двумерного газа в гетеропереходе по мере изменения двумерной электронной плотности от $n \approx 0,6 \cdot 10^{11} \text{ см}^{-2}$ (верхняя кривая) до $n \approx 6 \cdot 10^{11} \text{ см}^{-2}$ (нижняя кривая) [11]

центрации $n \approx 0,6 \cdot 10^{11} \text{ см}^{-2}$, а при $n \approx 6 \cdot 10^{11} \text{ см}^{-2}$ имеет место лишь слабый логарифмический рост ρ с понижением T . Хотя возможности для детального сравнения в этом эксперименте были весьма невелики, качественно получилось именно то, чего следует ожидать на основании скейлинговой гипотезы.

Более детальное сравнение удалось произвести на пленках Cu, Ag и Au с толщинами a в интервале от 0,3 до 2 нм.

Заметьте. Размерность образца определяется сравнением толщины a с неупругой длиной L_φ , поэтому эти пленки действительно двумерные объекты. А при достаточно низких температурах двумерными могут стать и гораздо более толстые пленки.

Результаты измерений на всех трех металлах построены в виде серии кривых в правой части рис. 6.10, а; при этом первоначальном построении в качестве T_0 использовалась величина 1 К, а величина y_{00} была выбрана равной $y_{00} = (2\pi)^{-1}(e^2/\hbar)$. Оставив y/y_{00} фиксированным, все экспериментальные точки удалось положить на одну кривую, сделав T_0 свободным параметром, своим для каждой из экспериментальных кривых. Для каждой из кривых на рис. 6.10, а это означает параллельный перенос вдоль горизонтальной оси на расстояние $-\ln T_0$. Хотя измерения на каждом отдельном образце укладываются лишь на небольшой участок получившейся сводной кривой, вся кривая перекрывает больше трех порядков по T/T_0 и содержит как участок слабой локализации (часть кривой $\ln y/y_{00} > 0$, рис. 6.10, б), так и активационный участок ($\ln y/y_{00} < -2$, рис. 6.10, в).

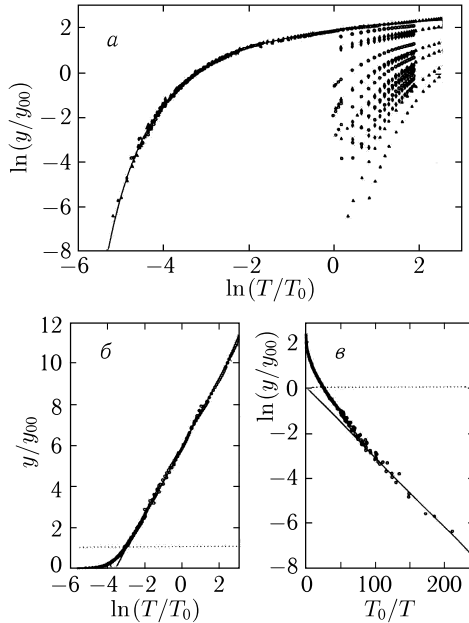


Рис. 6.10. а) Температурные зависимости $\ln y/y_{00}$ пленок из разных металлов и различной толщины (кривые справа) при помощи одного свободного параметра T_0 , т.е. при помощи параллельного горизонтального переноса, все ложатся на одну универсальную кривую. Асимптота этой кривой при больших y — это логарифмическая зависимость, характерная для слабой локализации (б), а при малых y это активационный закон Аррениуса (в) [12]

На активационном участке, т.е. в области сильной локализации, теория прыжковой проводимости допускает различные степени в формуле $y \propto \exp[-(T_0/T)^\nu]$. Согласно рис. 6.10, в, в данном эксперименте температурная зависимость $y(T)$ имеет степень $\nu = 1$, что указывает на существование жесткой щели шириной T_0 .

Сравните. На ультратонких пленках Ве наблюдаются степень $\nu = 1/2$ (см. рис. 4.7) и кулоновская щель (см. рис. 3.10 и рис. Б.5). Даже в виде ультратонких пленок материалы сохраняют индивидуальность.

Наблюдать кроссовер от слабой локализации к сильной удалось и на одном образце. На рис. 6.11 приведены измерения сопротивления в интервале от комнатной температуры до 0,2 К (больше трех порядков) на образце GaAs, δ -легированном кремнием, имеющем квазидвумерную конфигурацию (длинная полоска шириной 500 Å). Концентрация носителей в δ -слое определяется исходным уровнем легирования и не может быть изменена в процессе эксперимента. Как и в тонких металлических пленках, здесь реализуется принцип «один образец —

одна концентрация». Результаты измерения ниже 1 К укладываются на активационную зависимость $R \propto \exp(T_0/T)$ с $T_0 \approx 2,6$ К, а выше T_0 кривая $R(T)$ описывается формулами теории слабой локализации. В какой мере существенна квазиодномерность образца, не совсем ясно. Возможно, что использованные геометрические размеры оказались особо благоприятны для расширения температурного интервала измерений. Во всяком случае линейная зависимость сопротивления от $\ln T$ в области температур выше T_0 характерна именно для двумерных образцов.

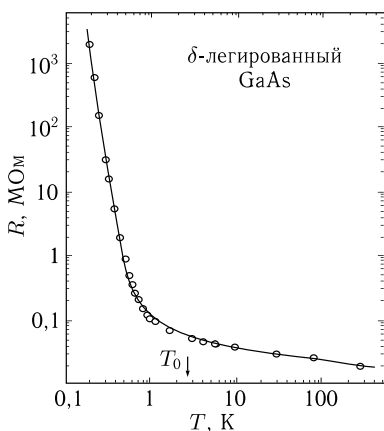


Рис. 6.11. Температурная зависимость сопротивления полоски шириной 0,05 микрон из δ -легированного GaAs [13]

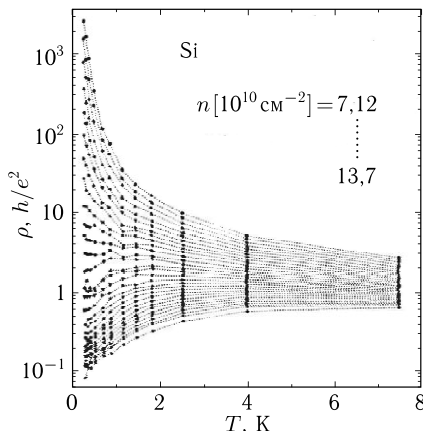


Рис. 6.12. Температурная зависимость сопротивления инверсионного слоя на поверхности Si при различной плотности носителей в слое [14]

Описанную гармонию между скейлинговой гипотезой и экспериментами нарушают результаты опытов на инверсионных слоях в кремнии, где концентрация регулируется напряжением на затворе. На образцах с рекордно большой подвижностью носителей в слое увеличение концентрации носителей n_2 в 2D-слое меняет знак производной $\partial\rho/\partial T$, так что при $n_2 \approx 1,4 \cdot 10^{11} \text{ см}^{-2}$ сопротивление падает в несколько раз по мере уменьшения температуры от 4 К до нескольких десятков мК (см. рис. 6.12). В этом существенное отличие от кривых для гетероструктуры GaAs/ $\text{Al}_x\text{Ga}_{1-x}\text{As}$ (на рис. 6.9) и существенное противоречие со скейлинговой гипотезой.

Интерпретация эволюции производной $\partial\rho/\partial T$ с концентрацией в инверсионных слоях в кремнии вызвала много споров. Однако, сейчас уже общепризнано, что благодаря электрон-электронному взаимодействию переход металл-изолятор в двумерном кремнии действительно происходит. Мы кратко вернемся к этому вопросу в последнем параграфе этой главы.

6.6. Квазиодномерные (q1D) системы

В параграфе 5.6 было показано, что в простейшей модели одномерного одноканального проводника со случайно расположенными одинаковыми рассеивателями происходит локализация при увеличении его длины L , причем эффективная длина свободного пробега l в достаточно коротком проводнике превращается в длину локализации ξ при увеличении длины, $\xi = l$. Требование об одинаковости рассеивателей не принципиально. Результат (5.45) для одноканального проводника подтверждается при помощи процедуры усреднения и в модели с разными рассеивателями, хотя эту процедуру следует проводить очень аккуратно (см. замечания после формул (5.36) и (5.46) и ссылки [9, 11] в гл. 5).

Запишем формулы для одноканального проводника в виде, который сделает их удобными для использования в сочетании со скейлинговой гипотезой. Для этого используем безразмерный кондактанс $y = Y/(e^2/2\pi\hbar)$ и соответствующее безразмерное сопротивление $\varrho = y^{-1}$. Согласно формуле (5.42), кондактанс y для «одноканальной проволоки» в металлическом пределе (т. е. не очень длинной) равен

$$y = l/L. \quad (6.35)$$

Применим формулу (5.31) для сопротивления дефекта в одноканальном проводнике, $\varrho = \mathcal{R}/T = 1/T - 1$, взяв в качестве дефекта цепочку из N барьеров, рассмотренную в параграфе 5.6. Согласно формуле (5.43), в локализационном пределе $T_N = se^{\alpha N} = se^{-L/\xi} = se^{-L/l}$ и

$$\varrho(L) = y^{-1}(L) = T^{-N} - 1 = \frac{1}{s}e^{L/l} - 1. \quad (6.36)$$

Когда цепочка барьеров короткая, так что $L/l \ll 1$, экспоненту в уравнении (6.36) можно разложить в ряд. Тогда из сравнения с выражением (6.35) для сопротивления цепочки барьеров в металлическом пределе получаем, что $s = 1$. Окончательно

$$y^{-1}(L) = e^{L/l} - 1. \quad (6.37)$$

Скейлинговое уравнение (6.37) определяет кондактанс $y(L)$ проволоки любой длины через саму длину L и «длину пробега в металлическом пределе» l . При помощи дифференцирования это уравнение можно привести к виду (6.3), получив тем самым кривую $\beta_1(y)$ на рис. 6.1 для размерности $d = 1$. Конечно, она получена из анализа одного частного вида 1D-системы, однако, скейлинговая гипотеза утверждает, что эта кривая является универсальной.

На границе $L = l = \xi$ между металлическим и локализационным режимами кондактанс одноканальной проволоки y принимает значение

$$y_\xi = (e - 1)^{-1} \approx 0,582, \quad \text{а} \quad R_\xi \approx 44,4 \text{ k}\Omega. \quad (6.38)$$

Воспользуемся универсальностью скейлинговой кривой и применим ее к многоканальным одномерным системам, обсуждавшимся в параграфе 5.5. Пусть 1D-система имеет ν каналов (размерно квантованных подзон под уровнем Ферми) — такую систему обычно называют квазиодномерной (q1D) проволокой. Квазиодномерные системы бывают двух видов. Во-первых, это может быть тонкая проволока квадратного или круглого сечения с характерным размером b , в несколько раз превосходящим π/k_F . Формула (5.22) позволяет убедиться, что при достаточно большой площади $S \approx b^2$ сечения проволоки число каналов $\nu \approx 2b^2/(\pi k_F)^2 \propto S$ (см. также формулы (5.22) в гл. 5 и (9.10) в гл. 9). Во-вторых, это может быть полоска двумерного электронного газа. В этом случае число каналов определяется шириной полоски, $\nu \propto b$. Такая q1D-система использовалась, например, в экспериментах, описанных в гл. 5 (см. там рис. 5.14 и 5.15). Мы будем называть проволоками оба типа q1D-систем.

Рассмотрим q1D-проволоку с ν каналами при $T = 0$. Пусть ее длина $L < \xi$, так что проволока находится в металлическом режиме и ее транспорт характеризуется длиной пробега l . Согласно формуле (6.35), кондактанс каждого канала равен $y_1 = l/L$, а кондактанс всей проволоки

$$y = \nu y_1 = \nu l/L. \quad (6.39)$$

При этом не имеет значения, что общая для всех каналов величина l сама по себе может контролироваться межканальным рассеянием.

Ввиду универсальности скейлинговой кривой (6.37), значение $y_\xi \approx 0,582$ является критическим не только для одноканальной проволоки, а для любой q1D-проволоки при $y \approx y_\xi$ реализуется равенство $L = \xi$ (см. рис. 6.13). Отсюда получаем обобщение соотношения (5.45)

$$\xi = \nu l. \quad (6.40)$$

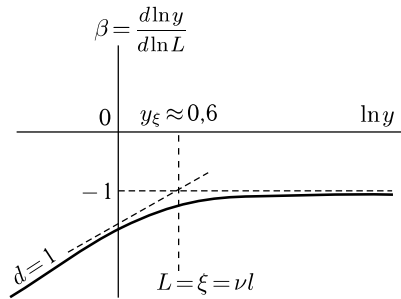


Рис. 6.13. Универсальная скейлинговая кривая для размерности $d = 1$

Представим себе, что мы имеем возможность измерять при разных (в том числе достаточно низких) температурах сопротивление отрезков различной длины L реальной тонкой проволоки. Пусть проволока статистически однородна, т.е. сопротивление отрезка однозначно определяется его длиной: $R = R(L)$. Каждый образец характеризуется тремя длинами: геометрической длиной L , корреляционной длиной ξ , определяемой числом и взаимным расположением примесей и других статических дефектов, и диффузионной длиной L_φ , которая зависит от температуры T . Пусть сначала T близка к нулю, так что $L_\varphi \gg \xi, L$.

Пока длина проволоки L невелика, так что ее кондактанс $y \gtrsim y_\xi$, и соответственно, сопротивление $R \lesssim 45 \text{ k}\Omega$, изображающая точка находится на правой асимптотике $\beta_1(y) = -1$ скейлинговой кривой. В этой «металлической» области справедлив закон Ома, т. е. $R \propto L$. Однако, при измерении более длинных отрезков той же проволоки $L > \xi$, когда изображающая точка попадет в локализационную область, а сопротивление станет $R \gtrsim 45 \text{ k}\Omega$, должен начаться экспоненциальный рост сопротивления при росте L . Необходимым условием локализации является выполнение неравенств

$$L_\varphi > L > \xi.$$

Если поднять температуру и тем уменьшить длину L_φ , сделав $L_\varphi < \xi$, то все волновые интерференционные процессы будут затухать на длине L_φ , а длинные проволоки будут вести себя классически: $R \propto L$ даже при очень больших L .

Пользуясь скейлинговой гипотезой, формально мы можем определить длину локализации ξ , имея лишь короткий кусок проволоки длиной $L \ll \xi$:

$$\xi \approx L \frac{45 \text{ k}\Omega}{R_0(L)},$$

где $R_0(L) = R(L, T \rightarrow 0)$ — значение сопротивления этого образца при $T = 0$, полученное экстраполяцией измеренной температурной зависимости.

Транспорт на расстояния $L > \xi$ в локализационной области обязательно включает квантовые переходы (прыжки), и именно вероятность таких прыжков контролирует проводимость. Это означает, что проводимость отрезка проволоки, принадлежащего локализационной области, следует описывать формулами прыжковой проводимости.

В проволоках электронный спектр происходит не из примесных уровней в запрещенной зоне, а из спектра свободных носителей под влиянием размерного квантования и случайного потенциала. Поэтому он не имеет вид узкой полосы с максимумом примерно посередине (ср. гл. 3). Его плотность состояний в широкой окрестности Ферми-уровня представляет собой монотонную медленно меняющуюся функцию. Поэтому естественно было бы ожидать, что в проволоках доминирующей является проводимость с переменной длиной прыжка. Действительно, на рис. 5.15 была продемонстрирована характерная для этого типа прыжковой проводимости зависимость $\ln Y \propto T^{-1/2}$ с переходом на $\ln Y \propto T^{-1/3}$ при увеличении ширины полосы b . Этот переход показывает, что в образцах реализуется моттовский вариант прыжков с переменной длиной, без кулоновской щели на Ферми-уровне.

С другой стороны, на рис. 6.11 ясно видно, что в q1D-проволоке, изготовленной на базе δ -легированного GaAs, в локализационной области имеет место активационный закон изменения сопротивления

$$R(T) \propto \exp(T_0/T). \quad (6.41)$$

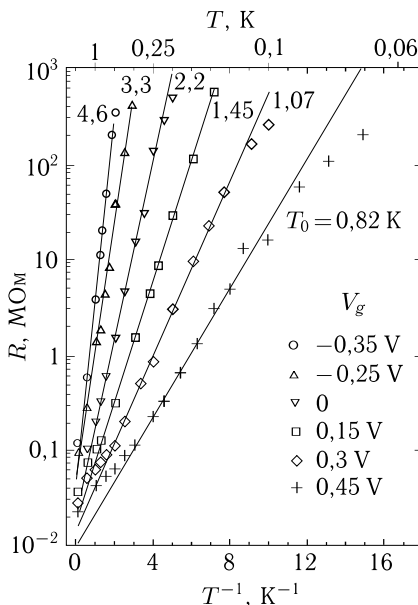


Рис. 6.14. Температурная зависимость сопротивления q1D-провода на базе δ -легированного GaAs при разных напряжениях V_g на затворе, т. е. при разной концентрации электронов в проволок [13]. Около кривых указана энергия активации T_0 , входящая в формулу (6.41)

Активационная зависимость (6.41) сохраняется и при изменении концентрации электронов в q1D-провонок на базе δ -легированного GaAs (см. рис. 6.14), хотя для самой большой концентрации ($V_g = 0,45$ В), когда температурную зависимость удается промерить до 60 мК, уже не совсем ясно, какую формулу, (6.41) или (4.24), следует предпочесть для обработки результатов.

Активационная зависимость (6.41) наблюдалась и на q1D-проводах другого происхождения. Например, она была подробно исследована на проволоках, изготовленных на базе инверсных слоев на поверхности кремния [15]. Появление чисто активационной зависимости (6.41) в сопротивлении проволок требует специального объяснения. При одномерном движении электрон не имеет возможности обогнуть какие-либо места с расположением центров, невыгодным для прыжков на дальние расстояния. Именно эти места, где электрон вынужден прыгать на один из дальних центров, определяют полное сопротивление всей проволоки. Мы уже сталкивались с этой особенностью одномерного движения — отсутствием обходных путей, когда обсуждали хаотические осцилляции в квазиодномерном канале полевого транзистора (рис. 5.14).

Именно отсутствие таких особых мест имелось в виду, когда несколькими абзацами выше мы вводили понятие о «статистиче-

ской однородности» проволоки. В экспериментах, представленных на рис. 6.14, для того, чтобы убрать хаотические осцилляции, измеряли 470 проволок, включенных в параллель. Поэтому энергия T_0 в наблюдавшемся в работе [13] активационном законе (6.41) определяется средним расстоянием $\delta\varepsilon$ между уровнями вблизи энергии Ферми

$$T_0 \approx \delta\varepsilon = \frac{1}{g_F S\xi}, \quad (6.42)$$

где g_F — плотность состояний на уровне Ферми, а $S\xi$ — локализационный объем. Отсюда следует второй способ измерения длины локализации ξ в q1D-проводах — по показателю экспоненты в температурной зависимости сопротивления в локализационном режиме, с использованием формулы (6.42). Этим способом, в частности, было проанализировано отрицательное магнетосопротивление в локализационной области у q1D-провода из δ -легированного GaAs [13]. Было показано, что оно обусловлено увеличением в магнитном поле длины локализации ξ .

6.7. Скейлинг и взаимодействия

При обосновании скейлинговой гипотезы в параграфе 6.1 предполагалось, что волновая функция каждого локализованного электрона независима от всех остальных и что она затухает на больших расстояниях как $\exp(-r/a_B)$, из чего следовала асимптотика (6.5). Это предположение может нарушаться при наличии электронных взаимодействий, а также в сильном магнитном поле. В этом параграфе мы постараемся хотя бы кратко упомянуть об известных случаях, когда взаимодействия нарушают соотношения, сформулированные в этой главе.

Начнем со спин-орбитального взаимодействия. Его наличие и связанная с ним антилокализация усложняют вопрос об универсальности кривых $\beta(\ln y)$ на рис. 6.1. Ведь спин-орбитальное взаимодействие меняет знак квантовой поправки и, согласно таблице в начале этой главы, должен измениться знак производной в правой части кривых $\beta(\ln y)$. Это особенно существенно в двумерном случае, при $d = 2$, когда квантовая поправка (2.25) имеет вид

$$\Delta\sigma_2 \approx -(e^2/\hbar) \int_{\tau}^{\tau_\varphi} \frac{dt}{t} \left(\frac{3}{2} e^{-t/\tau_{so}} - \frac{1}{2} \right). \quad (6.43)$$

Может показаться, что учет спин-орбитального взаимодействия делает утверждение об отсутствии перехода металл–изолятор в двумерной системе принципиально неверным. Действительно, поскольку речь идет о проводимости при $T = 0$, положим в интеграле (6.43) верхний предел $\tau_\varphi = \infty$. В то же время τ_{so} остается конечным, потому что

вероятность переворота спина при упругом рассеянии не зависит от температуры. Множитель

$$\frac{3}{2} e^{-t/\tau_{so}} - \frac{1}{2\tau_{so}},$$

стоящий в интеграле (6.43) в скобках, в области $t > \tau_{so}$ становится порядка $-\frac{1}{2}$. Это делает интеграл, расходящийся при больших t , отрицательным, а квантовую поправку к проводимости положительной. Кривая $\beta(\ln y)$ приобретает максимум, расположенный в верхней полуплоскости $\beta > 0$ и критическую точку y_c . На рис. 6.15 эта кривая изображена пунктиром. Сплошной линией нарисована кривая, построенная без учета спин-орбитального взаимодействия в предположении $\tau_{so} = \infty$.

Однако переход к $T = 0$, т. е. переход $\tau_\varphi \rightarrow \infty$ в верхнем пределе интеграла (6.43), требует некоторой осторожности. Это видно из следующих рассуждений.

Пусть классический кондактанс образца равен y_0 . Начнем вычислять квантовую поправку к проводимости, постепенно увеличивая верхний предел τ^* в интеграле (6.43), начиная с τ . Это означает, что мы начинаем с учета интерференции на самых маленьких диффузионных контурах, а затем, мысленно увеличивая размер образца L или диффузионную длину L_φ , постепенно учитываем контуры со все большим числом звеньев. Пока $\tau^* < \tau_{so}$, поправка к проводимости растет по абсолютной величине, оставаясь отрицательной. Мы двигаемся в сторону уменьшения $\ln y$, причем это движение происходит вдоль нижней (сплошной) кривой. Когда $\tau^* = \tau(L_\varphi/l)^2$ достигнет τ_{so} , начнет развиваться антилокализация и изображающая точка перепрыгнет на верхнюю (пунктирную) кривую. Весь вопрос в том, когда это произойдет. Чем сильнее спин-орбитальное взаимодействие, тем меньше τ_{so} и тем правее будет точка перескока. Перескок в точке $y_1 > y_c$ означает, что при увеличении размеров пленки изображающая точка будет двигаться вдоль пунктирной кривой направо, кондактанс будет расти, и пленка останется в металлическом состоянии. При перескоке в точке $y_2 < y_c$ изображающая точка останется в нижней полуплоскости и будет продолжать двигаться влево в состояние изолятора.

Для того чтобы оценить критическое значение τ_{so} , заметим, что критическая точка на пунктирной кривой расположена примерно там

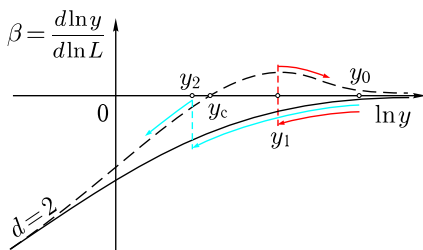


Рис. 6.15. Два варианта кривой $\beta_2(\ln y)$ для размерности $d = 2$

же, где и кроссовер от логарифмической к экспоненциальной зависимости кондактанса от температуры на сплошной кривой. Используя формулу (6.32) и выражение (2.7) для L_φ , получим оценку значения τ_{so}^* , при котором точка y_1 попадает на критическое значение y_c :

$$k_F l = \frac{1}{2} \ln \tau_{so}^* / \tau, \quad \tau_{so}^* = \tau e^{2k_F l}. \quad (6.44)$$

Для образования металлического двумерного состояния требуется $\tau_{so} \ll \tau_{so}^*$. Поскольку по определению $\tau_{so} > \tau$, для τ_{so} возникает допустимый интервал

$$\tau \ll \tau_{so} \ll \tau \exp(2k_F l). \quad (6.45)$$

В сильно разупорядоченных пленках удовлетворить эти неравенства невозможно, потому что $k_F l \sim 1$. В таких пленках спин-орбитальное взаимодействие ничего существенно не меняет.

В пленках с $k_F l \gg 1$ вопрос имеет чисто академический характер, потому что температура кроссовера (6.34) нереально низкая. Однако в принципе хорошо проводящая пленка с большим $k_F l$ благодаря спин-орбитальному взаимодействию может остаться металлической при абсолютном нуле температуры.

В том, что для существования обсуждаемого перехода необходим беспорядок, и в то же время, согласно соотношениям (6.45), он может наблюдаться только в сверхчистых пленках, нет противоречия. Ведь исходное утверждение состоит в том, что в 2D-системе без взаимодействия переход изолятора в металлическое состояние отсутствует при любом беспорядке. Из соотношения (6.45) следует, что спин-орбитальное взаимодействие способно «перебороть» только очень слабый беспорядок.

Следует отметить, что и квантовая поправка от межэлектронной интерференции, которая обсуждалась в параграфе 2.4, тоже может изменять знак. Она становится положительной при наличии сверхпроводящего взаимодействия. Однако, сверхпроводящее взаимодействие оказывает и более существенное воздействие на эволюцию транспортных свойств материалов, в которых оно присутствует, приводя к переходам изолятор–сверхпроводник вместо переходов изолятор–нормальный металл. Мы ограничимся здесь только одним примером (см. рис. 6.16) и ссылкой на обзор [17], где эти переходы подробно обсуждаются.

Уравнение (6.3) устанавливает, что в системе имеется только один параметр y , определяющий расстояние Δy до перехода. Для того чтобы это уравнение осталось справедливым после включения взаимодействия, требуется, чтобы взаимодействие не зависело от размера системы L (а при конечной температуре T от длины боя фазы L_φ). Между тем, это требование часто нарушается. Примером может служить только что обсуждавшееся спин-орбитальное взаимодействие. Из рассуждений, сопровождавших рис. 6.15, и формул (2.7) и (6.44)

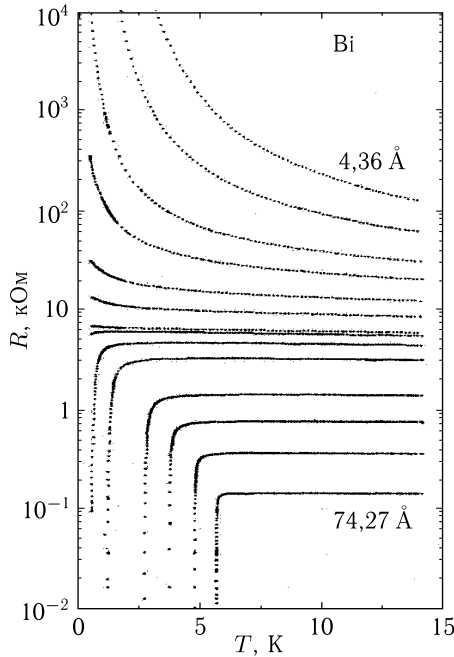


Рис. 6.16. Переход сверхпроводник–изолятор в пленке аморфного Bi, нанесенного на подложку из аморфного Ge толщиной 5 Å [16]

следует, что оно включается лишь при

$$L > L^* \approx l \sqrt{\tau_{so}^*/\tau} \approx l e^{k_F l}, \quad (6.46)$$

а при $L < L^*$ отсутствует.

При включении взаимодействия Θ , которое меняется при изменении L или T , состояние системы определяется двумя параметрами, y и Θ , а вместо одного уравнения (6.3) появляются два

$$\begin{aligned} \frac{d \ln y}{d \ln L} &= \beta_y(\ln y, \ln \Theta), \\ \frac{d \ln \Theta}{d \ln L} &= \beta_\Theta(\ln y, \ln \Theta), \end{aligned} \quad (6.47)$$

и скейлинг из однопараметрического становится двухпараметрическим.¹⁾ По сути дела, перескоки на рис. 6.15 появились из-за того, что мы попытались двухпараметрический скейлинг свести к однопараметрическому, прибегая в своих рассуждениях к двум вариантам

¹⁾ Более подробно о схеме двухпараметрического скейлинга можно прочитать в обзоре [17].

кривой $\beta_2(\ln y)$ вместо использования двухпараметрической потоковой диаграммы.

Анализ [18] в рамках двухпараметрического скейлинга показал, что межэлектронное взаимодействие в 2D-системах действительно может сделать непрерывный переход металл–изолятор возможным и что поведение двумерного электронного газа на поверхности кремния [14] определяется именно близостью к такому переходу. Подвижность электронов в кремнии в экспериментах [14] была выше, чем в других материалах [11, 12], где признаков такого перехода не было. Получается, что и здесь взаимодействие может привести к переходу металл–изолятор только при достаточно слабом беспорядке.

Последовательного рассмотрения двухпараметрического скейлинга в 3D-системах пока нет. Поэтому экспериментальные наблюдения об отсутствии локализации в 3D-системах с большой электронной плотностью (гл. 1) и о скачкообразном уменьшении проводимости до нуля в 3D-системе с большим беспорядком при уменьшении электронной плотности в области $n \gtrsim 10^{21} \text{ см}^{-3}$ (гл. 5, рис. 5.10) не имеют пока адекватного теоретического описания.

В начале этого параграфа упоминалось о том, что условия применимости однопараметрических скейлинговых кривых рис. 6.1 нарушаются и в сильном магнитном поле. Мы обсудим это в конце гл. 9, посвященной целочисленному квантовому эффекту Холла (рис. 9.22–9.27). Кроме того, в гл. 10 излагаются физические основы общей теории квантовых фазовых переходов, одним из частных случаев которых являются переходы металл–изолятор. Тогда некоторые из введенных здесь понятий (например, понятие критической области) приобретут более глубокий физический смысл.

6.8. Контрольные вопросы и задачи

1. На какую часть кривой $\beta_3(\ln y)$ попадут точки, отображающие высокорезистивные металлические сплавы, описанные в гл. 1 (например, те, которые подчиняются правилу Мооыа)?

На правую почти горизонтальную часть.

2. Написать в аналитическом виде формулу $\beta(\ln y)$ для одномерной системы.

3. Вывести соотношение между длиной свободного пробега l и длиной локализации ξ в многоканальном одномерном проводнике.

Формула (6.40): $\xi = \nu l$.

4. Одномерный проводник длиной L при температуре T имеет сопротивление около 50 кОм. При этом определяемая рассеянием на фононах неупругая длина $L_\varphi \approx L$. Каково будет сопротивление, если

а) длина проводника L останется прежней, а температура T уменьшится в 10 раз?

б) температура T останется прежней, а длина проводника L увеличится в 10 раз (при сохранении локальных значений параметров τ и τ_φ)?

в) длина проводника L увеличится в 10 раз, а температура T уменьшится в 10 раз?

- а) практически не изменится;
- б) увеличится примерно в 10 раз;
- в) увеличится экспоненциально.

Список литературы

1. *Mott N. V.* Metal-insulator transitions. — Taylor & Francis, 1990.
2. *Lee P. A., Ramakrishnan T. V.* Disordered electronic systems // *Rev. Modern Phys.* **57**, 287 (1985).
3. *Kramer B., MacKinnon A.* // *Rep. Prog. Phys.* **56**, 1469 (1993).
4. *Abrahams E., Anderson P. W., Licciardello D. C., Ramakrishnan T. W.* // *Phys. Rev. Lett.* **42**, 673 (1979).
5. *Shlimak I., Kaveh M., Ussyshkin R.* et al. // *Phys. Rev. Lett.* **77**, 1103 (1996); *J. Phys. Cond. Matt.* **9**, 9873 (1997).
6. *Shlimak I., Kaveh M., Ussyshkin R., Ginodman V., Resnik L.* // *Phys. Rev. B* **55**, 1303 (1997).
7. *Федер Е.* Фракталы. Мир, 1991 [Перевод книги *Feder J.* *Fractals*, Plenum Press, 1988].
8. *Huckestein B.* // *Rev. Mod. Phys.* **67**, 357 (1995).
9. *Evers F. and Mirlin A. D.* // *Rev. Mod. Phys.* **80**, 1355 (2008).
10. *Richardella A., Roushan P., Mack S., Zhou B., Huse D. A., Awschalon D. D., and Yazdani A.* // *Science* **327**, 665 (2010).
11. *Van Keuls F. W., Mathur H., Jiang H. W., Dahm A. J.* // *Phys. Rev. B* **56**, 13263 (1997).
12. *Hsu S.-Y., Valles J. M.* // *Phys. Rev. Lett.* **74**, 2331 (1995).
13. *Havin Yu., Gershenson M., Bogdanov A.* // *Phys. Rev. B* **58**, 8009 (1998).
14. *Kravchenko S. V., Mason W. E., Bowker G. E., Furnaux J. E., Pudalov V. M., Iorio M. D.* // *Phys. Rev. B* **51**, 7038 (1995).
15. *Kwasnick R. F., Kastner M. A., Melngallis J., and Lee P. A.* // *Phys. Rev. Lett.* **52**, 224 (1984).
16. *Haviland D. B., Liu Y., Goldman A. M.* // *Phys. Rev. Lett.* **62**, 2180 (1989).
17. *Гантмахер В. Ф., Долгополов В. Т.* // *УФН*, **180**, 3 (2010).
18. *Punnoose A. and Finkel'stein A.* // *Science* **310**, 289 (2005).

Глава 7

ХИМИЧЕСКАЯ ЛОКАЛИЗАЦИЯ

Зонная теория металлов с ее концепцией перекрытия энергетических зон не объясняет, а скорее описывает металлические свойства вещества. Фундаментальная причина существования металлического состояния состоит в том, что в изолированном металлическом атоме потенциальная яма для валентных электронов сравнительно неглубока, так что в конденсированном состоянии возмущение со стороны соседних металлических атомов приводит к делокализации валентных электронов. С такой точки зрения деление элементов на металлы и металлоиды обусловлено строением атомов; металлы сосредоточены в левом нижнем углу таблицы Менделеева, а граница между металлами и металлоидами размыта и весьма условна.

Поскольку валентные электроны в металлах делокализованы, электронная концентрация в них не меньше, чем концентрация атомов. Такого же порядка и электронная концентрация в сплавах. В начале гл. 1 материал с такой концентрацией был назван стандартным металлом. Там же был сформулирован вопрос: можно ли устроить переход металл–изолятор в веществе с большой электронной плотностью (см. также последний параграф и рис. 1.8 в гл. 1 и рис. 5.8 в гл. 5). И там же собран экспериментальный материал, из которого следует, что увеличение беспорядка, как статического (например в аморфных сплавах) так и динамического (рост температуры), к переходу не приводит. Создавая в сплаве максимальный беспорядок, мы лишь подводим его к крайней черте. Для того чтобы произошел переход, требуется дополнительно заменить часть металлических атомов на атомы металлоидов, уменьшив таким образом концентрацию n делокализованных электронов до критического значения.

Введение атомов металлоидов может оказаться вдвойне эффективно в том смысле, что концентрация металлических атомов не всегда однозначно определяет концентрацию n делокализованных или «потенциально делокализуемых» электронов. Если атомы металла и металлоида могут образовывать устойчивые химические молекулы, то металлические электроны уходят на химические связи; из неглубокой потенциальной ямы металлического атома они переходят в гораздо более глубокую яму молекулы и поэтому остаются локализованными, невидя на окружение этой молекулы. Поэтому эффективная электронная концентрация n , которая должна влиять на положение материала на

фазовой диаграмме металл–изолятор, уменьшается еще и потому, что возникают химические связи.

Концепция блокирования части «потенциально свободных» электронов химическими связями позволяет сформулировать вопрос: *Нельзя ли сконструировать глубокие потенциальные ямы, используя только металлеские атомы, что превратило бы в изолятор материал, в котором нет атомов металлоидов, и который должен был бы быть стандартным металлом?* Экспериментальный материал, собранный в этой главе, указывает на то, что это возможно.

7.1. Интерметаллические комплексы в двухкомпонентных расплавах ¹⁾

Уже давно известно, что удельное сопротивление жидкого расплава двух очень хороших металлов может меняться во много раз, даже на порядки, в зависимости от относительной концентрации двух компонентов, достигая максимума при некотором рациональном отношении атомных концентраций типа 1:1 или 1:3 или 1:4. Для получения подобных данных нужно уметь при фиксированной температуре измерять сопротивление в зависимости от концентрации компонентов сплава. В качестве примера кратко опишем экспериментальную установку, использовавшуюся для измерения системы Na–Pb (рис. 7.1).

Измерения производились обычным четырехконтактным способом в двух плечах U-образной трубки из нержавеющей стали диаметром около 4 мм, заполненной исследуемым расплавом. На обоих плечах трубки снаружи были припаяны по четыре медных провода для электрических измерений. Предварительно, перед заполнением расплавом, собственные сопротивления плеч были измерены в зависимости от температуры и могли быть учтены в виде поправок. Оба

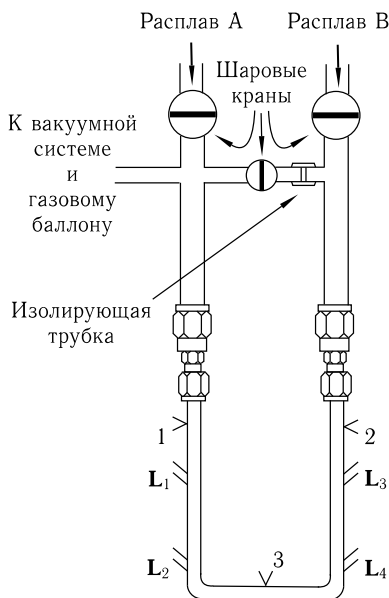


Рис. 7.1. Ячейка для измерений удельного сопротивления расплава. 1, 2 и 3 — термодары, L_1 и L_2 — медные провода для измерения сопротивления в левом колене, L_3 и L_4 — то же в правом колене [1, 2]

¹⁾ Экспериментальные подробности можно найти в обзорах [1, 2].

конца U-образной трубки при помощи гаечных разъемов были подсоединены к трубкам большего диаметра, которые служили резервуарами расплава. При помощи чистого гелия и откачки расплав можно было выдавливать то в один, то в другой резервуар и тем самым перемешивать. Перемешивание производилось до тех пор, пока равенство сопротивлений в обоих плечах трубки не подтверждало однородности расплава. Перед окончательными измерениями сопротивления давление газа и уровни расплава в обоих плечах выравнивались. После измерения сопротивления одного какого-то фиксированного состава в расплав добавлялось определенное количество одного из компонентов и вся процедура повторялась.

На рис. 7.2 приведены результаты измерений расплава Na–Pb при температуре 725 °С. Видно, что соотношение концентраций Na к Pb равное 4:1 явно выделено. Добавка 20% Pb увеличивает сопротивление по сравнению с чистым Na почти в 20 раз. При этом значение сопротивления в максимуме получилось больше максимально возможного сопротивления ρ^* стандартного металла

$$\rho^* \approx \frac{\hbar}{e^2} n^{-1/3} \approx (200 \div 300) \text{ мкОм} \cdot \text{см} \quad (7.1)$$

при длине пробега l порядка среднего расстояния между носителями $n^{-1/3}$ и примерно равной ему длины де-бройлевской волны k_F^{-1} (см. формулы (1.5), (1.6) из гл. 1). Примерно так же ведет себя и система Li–Pb.

Замена Li и Na на более тяжелые щелочные металлы K, Rb или Cs приводит к существенному изменению диаграммы сопротивления — концентрация. Для примера на рис. 7.3 приведена диаграмма для системы Rb–Pb. Положение максимума на диаграмме сместилось к рациональному отношению концентраций 1:1 компонентов Rb и Pb, а значение сопротивления в максимуме стало в несколько раз больше. Оно уже на порядок превышает максимальное сопротивление стандартного металла (1.5), (1.6). Очень важно, что малое отклонение от соотношения 1:1 приводит к резкому падению сопротивления. Как видно из обзорных графиков на рис. 7.4, точно так же ведут себя и расплавы щелочных металлов с другим четырехвалентным металлом — оловом. При этом в расплавах на основе Cs реализуются еще большие значения удельного сопротивления.

Большие значения сопротивления означают, что вблизи соответствующих отношений концентраций расплав перестает быть стандартным металлом в том смысле, что часть носителей в нем каким-то образом связана, а оставшаяся эффективная концентрация $n_{\text{eff}} \ll \ll 4 \cdot 10^{22} \text{ см}^{-3}$ (ср. с (1.4)). Действительно, для системы Rb–Pb величина $\rho \approx 2200 \text{ мкОм} \cdot \text{см}$ раз в 10 больше максимального значения для стандартного металла $\rho^* \approx (200 \div 300) \text{ мкОм} \cdot \text{см}$. Согласно формулам (1.5), (1.6) отсюда следует, что в расплаве число свободных носителей

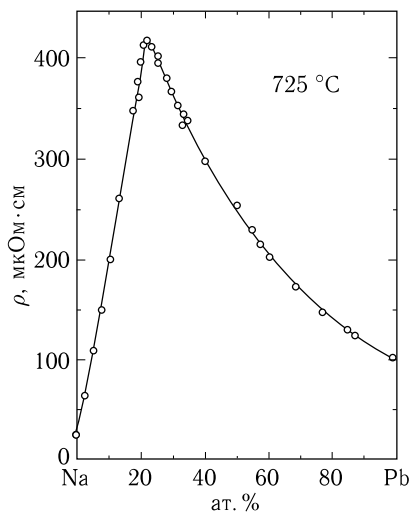


Рис. 7.2. Сопротивление расплавов системы Na–Pb при температуре 725 °С. Максимум достигается при концентрации свинца $C_{Pb} \approx 20\%$, когда возникают устойчивые конфигурации $PbNa_4$ [3]

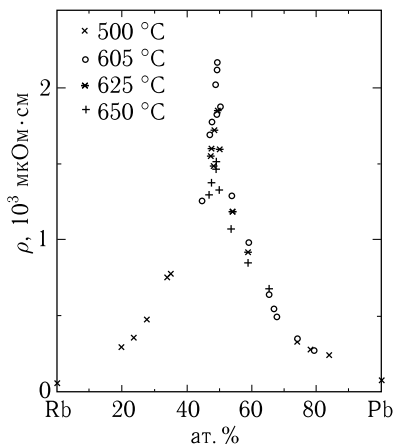


Рис. 7.3. Сопротивление расплавов системы Rb–Pb при различных температурах [4]. Максимум достигается при концентрации свинца $C_{Pb} = 50\%$; устойчивыми являются конфигурации Pb_4Rb_4

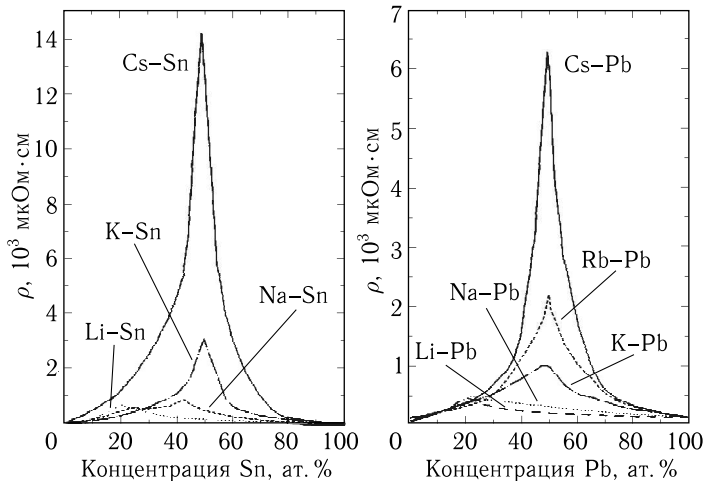


Рис. 7.4. Обзорные графики сопротивления расплавов систем Sn–В и Pb–В (В — щелочной металл). Сопротивление сплавов с Li и сплава Pb–Na имеет максимум при $C_A = 20\%$ ($A = Pb, Sn$); сопротивление сплава Sn–Na имеет два максимума при $C_{Sn} \approx 25\%$ и 45% ; все остальные — при $C_A = 50\%$ [5]

не больше, чем $10^{-2} \div 10^{-3}$ от обычного числа носителей в стандартном металле.

Из рациональности соотношений (1:4, 1:1) концентраций компонентов следует, что рост сопротивления связан с образованием комплексов. Положение максимума в расплавах на основе Li и Na с несомненностью указывает на существование в расплаве комплексов Na_4Pb и Li_4Pb . В них оказывается запертой основная часть электронов. На валентных оболочках пяти входящих в такой комплекс атомов имеется восемь электронов. По-видимому они группируются в одну устойчивую внешнюю оболочку иона Pb^{4-} , а четыре щелочных иона располагаются вокруг, удерживаемые кулоновскими силами. Они создают барьер, благодаря которому восемь электронов на внешней оболочке Pb удерживаются внутри этой электрически нейтральной атомной конфигурации и не участвуют в проводимости (рис. 7.5, а).

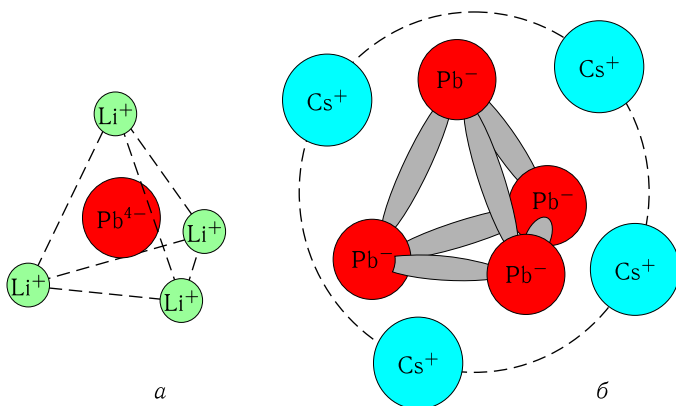
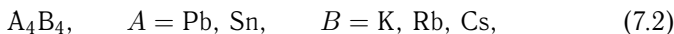


Рис. 7.5. а) Ионные конфигурации из атомов Pb или Sn и легких щелочных металлов Li или Na; б) то же с участием атомов тяжелых щелочных металлов K, Rb или Cs

Для щелочных атомов большего размера такая конфигурация становится менее выгодной, потому что из-за взаимного отталкивания ионы щелочного металла не могут расположиться близко от иона свинца. Это приводит к качественному изменению образующихся конфигураций. Эти структуры известны. По имени немецкого химика, предложившего в 30-х годах принцип формирования ионных конфигураций, они называются структурными единицами Зинтля. Если электрон перейдет с щелочного атома на атом Pb, то у иона Pb^- на внешней оболочке окажется пять электронов, как у атома P или As. Последние, как известно, образуют в газовой фазе тетраэдральные молекулы P_4 или As_4 . При этом вблизи каждого атома оказывается по восемь электронов: пять своих и по одному из ковалентной связи с каждым из трех соседей в тетраэдре. Ионы Pb^- образуют такие же

тетраэдры, а суммарный электрический заряд $-4e$ такого тетраэдра компенсируется примыкающими к нему четырьмя ионами щелочного металла (рис. 7.5, б). Аналогично ведут себя ионы Sn^- , которые образуют тетраэдры $(\text{Sn}_4)^{4-}$, обрамленные четырьмя щелочными ионами.

Способность таких комплексов служить электронными ловушками выше, чем у комплексов Na_4Pb и Li_4Pb . В каждой структурной единице



оказываются запертыми 20 валентных электронов.

Аналогично ведут себя бинарные расплавы щелочных металлов с некоторыми другими металлами. Абсолютным чемпионом с точки зрения образования эффективных электронных ловушек является изотомный расплав двух совершенно идеальных металлов: щелочного Cs и благородного Au. Как видно из рис. 7.6, образование комплексов в расплаве уменьшает проводимость на 4 порядка. Она становится сравнимой с проводимостью расплава солей ($3 \text{ Ом}^{-1} \times \text{см}^{-1}$ у CsAu и $1 \text{ Ом}^{-1} \cdot \text{см}^{-1}$ у расплава соли CsCl).

Вопросы о том, где, когда, какие и сколько конфигураций Зинтля образуются в металлическом расплаве и каковы энергии связи этих конфигураций, относятся к компетенции квантовой химии и химической термодинамики. Однако, результат можно прокомментировать на физическом языке. Говоря об отдельном атоме Cs, мы полагаем, что из его 55 электронов 54 входят в электронный остов и обычно его не покидают, а в формировании металлических свойств материала участвует

только один валентный электрон. Соответственно, атом Pb выдает в общее пользование не более 4 валентных электронов, а остальные 78 всегда находятся внутри иона. Каждая конфигурация A_4B_4 из (7.2) является почти сферической ямой для $N = 4 \cdot 4 + 4 = 20$ валентных электронов, которые располагаются внутри нее на последовательных уровнях энергии. Покинуть конфигурацию могут только один-два электрона с верхних уровней. Другими словами, из 20 бывших валентных электронов 18–19 уходят в ионный остов конфигурации, т.е. перестают быть валентными. Поэтому вдоль шкалы абсцисс на рис. 7.4 можно поставить новые электронные концентрации n_{eff} : на левом конце

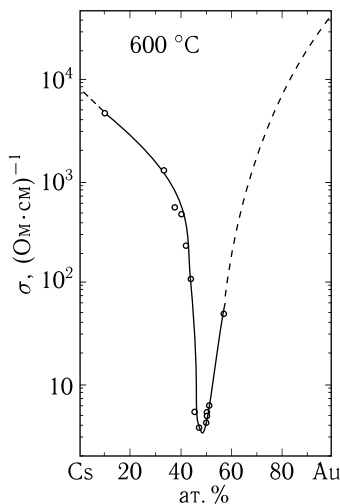


Рис. 7.6. Проводимость расплавов системы Cs–Au при температуре 600 °C [1, 6]

1 эл./атом, на правом конце 4 эл./атом, в середине при концентрации $C_A = 50\%$ находится минимум $n_{\text{eff}} \approx 0,12\text{--}0,25$ эл./атом.

Вопрос в том, хватит ли этого уменьшения n для того, чтобы электроны локализовались при низкой температуре. Чтобы формализовать этот вопрос, введем концентрацию конфигураций Зинтля $\mathcal{N}$ и используем модель структурного беспорядка (параграф 5.2), в которой случайный потенциал имеет вид

$$V(\mathbf{r}) = \sum_{\mathbf{R}_i} v(\mathbf{r} - \mathbf{R}_i), \quad 1 \leq i \leq \mathcal{N}. \quad (7.3)$$

Здесь $v(\mathbf{r})$ это потенциал отдельной ионной конфигурации, заряженной до $+e$, а беспорядок определяется хаотичностью множества векторов $\mathbf{R}_i$, задающих расположение конфигураций. Для определенности мы полагаем, что каждая конфигурация может выдать в общее пользование только один электрон. Локализация определяется параметром $a_B n_{\text{eff}}^{1/3}$, где a_B — длина затухания волновой функции единственного валентного электрона конфигурации, а $n_{\text{eff}} = n - (N - 1)\mathcal{N}$. Когда $C_A = 50\%$, все атомы объединены в конфигурации, а $n_{\text{eff}} = \mathcal{N} = n/N$. При малом параметре

$$a_B n_{\text{eff}}^{1/3} \ll 1 \quad (7.4)$$

эта модель описывает изолятор: каждый электрон локализован в своей яме, а вне ямы его волновая функция затухает на расстоянии a_B .

Сравните. При описании моноэлементных жидких металлов тот же потенциал (7.3) рассматривается как рассеивающий, так как предполагается, что вместо (7.4) справедливо обратное неравенство, а электроны полностью делокализованы.

В рамках такого описания можно качественно объяснить температурную зависимость проводимости CsAu, представленную на рис. 7.7.

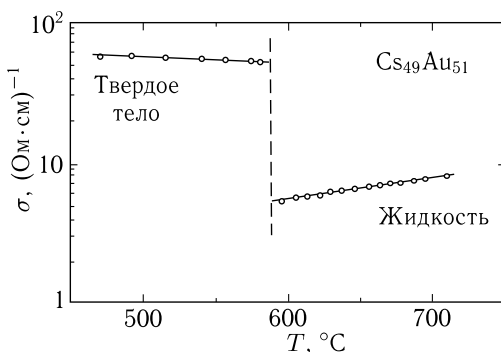


Рис. 7.7. Проводимость сплава Cs–Au с 51% Au как функция температуры в жидком и твердом состояниях [1, 7]

Изменение с температурой проводимости расплава CsAu в жидком состоянии определяется изменением числа термически возбужденных носителей; поэтому выше температуры плавления проводимость растет с температурой. При кристаллизации появляются кластеры резонансных ям, имеющих одинаковых и одинаково расположенных соседей. Хаотичность расположения ям частично сохраняется только в меру нестехиометрии и наличия кристаллических дефектов и межкуристаллитных границ. Следствием роста числа резонансных ям является частичная делокализация и десятикратный рост проводимости.

Однако заметьте. Проводимость кристаллического CsAu на рис. 7.7 все равно раз в 50 меньше, чем предельная проводимость $1/\rho^*$ стандартного металла (7.1). При этом неясно, в какой мере и как она зависит от отклонений от стехиометрии, количества дефектов, температуры и т. д.

Поскольку расплавы по определению находятся при высокой температуре, кривые на рис. 7.3 или 7.6 не следует считать демонстрациями переходов металл–изолятор. Для того чтобы ввести эти материалы в круг объектов, описываемых теорией переходов металл–изолятор, их следовало бы закалить, превратив в стекло. Тогда, например, из низкотемпературных зависимостей транспортных характеристик вероятно можно было бы получить количественные параметры электронных ловушек. На сегодняшний день попыток закалки таких расплавов с целью исследования их низкотемпературных свойств не известно.

7.2. Квазикристаллы ¹⁾

Введение трансляционной симметрии — это не единственный способ установления дальних корреляций на множестве векторов $\mathbf{R}_i$ в модели структурного беспорядка (7.3). Другим способом является установление квазикристаллического дальнего порядка.

Трансляционная симметрия, всегда имеющаяся в кристаллах, допускает существование осей симметрии только 2, 3, 4 и 6 порядков. Вместе с тем нетрудно себе представить, что из всевозможных локальных конфигураций небольшого количества атомов $A_n B_m C_p$ определенных элементов А, В и С конфигурация с наименьшей энергией может иметь иные оси симметрии, например оси 5 порядка. Тогда обычный кристалл из материала состава $A_n B_m C_p$ можно образовать двумя способами. Можно пожертвовать оптимальной локальной симметрией, пожертвовав в связи с этим некоторой энергией, и образовать кристалл с иной конфигурацией ближайших соседей у каждого

¹⁾ Подробное обсуждение различных аспектов физики квазикристаллов можно найти, например, в [8].

атома, но с трансляционной симметрией. Существует, однако, и иная возможность. Расположим оптимальные конфигурации из $n + m + p$ атомов в узлах какой-либо кристаллической решетки, например, объемноцентрированного куба. Тогда проигрыш в энергии возникнет за счет рассогласования и искажений на стыках этих конфигураций, где

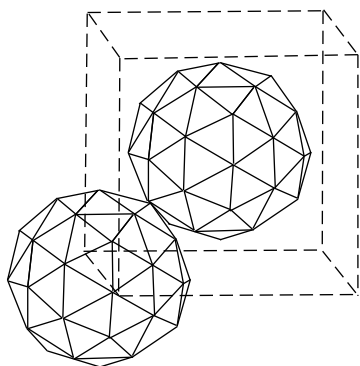


Рис. 7.8. Кристаллическая структура сплава $\alpha(\text{AlMnSi})$. Каждый икосаэдр Макея содержит более 50 атомов [9]

ближний порядок окажется заведомо не оптимальным. Тем не менее у некоторых веществ существуют именно такие кристаллические структуры. Они называются кристаллическими аппроксимантами или прототипами квазикристаллов. Именно так устроен кристаллический сплав $\alpha(\text{AlMnSi})$, оптимальные конфигурации в котором называются икосаэдрами Макея (рис. 7.8).

Оказывается, однако, что можно вообще отказаться от трансляционной симметрии, плотно заполняя пространство оптимальными конфигурациями. Принципиально такую возможность в двумерном пространстве демонстрирует покрытие Пенроуза

в нижней части рис. 7.9: плоскость плотно, без пропусков и перекрытий, покрывается ромбами двух типов с острыми углами $2\pi/5$ и $\pi/5$. Эти ромбы изображены на рисунке вверху слева. Поскольку при покрытии они примыкают друг к другу заранее помеченными вершинами, правильно заданные функции на ромбах $F(\mathbf{r})$ при стыковке ромбов не испытывают разрывов, образуя непрерывную аперiodическую функцию, отдельные участки которой повторяются на плоскости бесконечное количество раз. Сверху справа на рис. 7.9 показано отдельно расположение вершин ромбов (в масштабе 1:2). Из-за отсутствия трансляционной симметрии довольно трудно усмотреть наличие корреляций. Однако дальний порядок в системе имеется: ромбы располагаются на плоскости вполне определенным, хотя и неоднозначным способом.

Квазикристаллы устроены по этому же принципу. Оптимальные конфигурации с осями симметрии высокого порядка, как ромбы Пенроуза, выстраиваются в трехмерную сеть, не имеющую трансляционной симметрии, но обладающую дальним порядком. Сейчас уже известно много семейств таких материалов. Большинство из них являются металлическими сплавами в том смысле, что состоят только из металлических атомов: Al–Mn, Ga–Mg–Zn, Al–Cu–Fe, Al–Pd–Re и т. д. Базисные конфигурации при этом могут быть довольно сложными. Например в квазикристаллах состава Al–Pd–Mn они состоят из трех

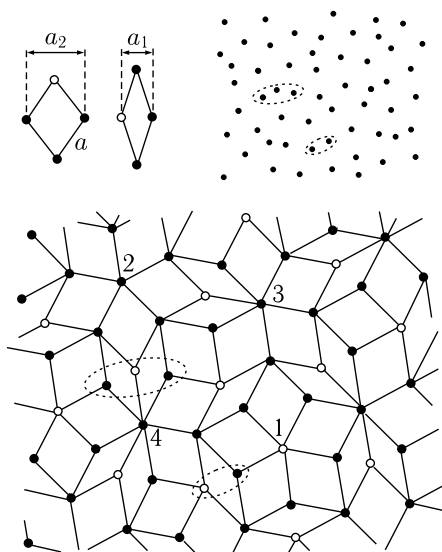


Рис. 7.9. Покрытие Пенроуза. Внизу: покрытие плоскости без зазоров и перекрытий двумя сортами ромбов, изображенныхверху слева: с одинаковыми сторонами a и с острыми углами, соответственно, $2\pi/5$ и $\pi/5$ (вершины, помеченные полыми кружками, в покрытии примыкают друг к другу). Вверху справа: множество вершин ромбов изображенного внизу фрагмента покрытия в масштабе 1:2. Хотя узлы 1–4 каждый имеет ближайших соседей только на расстояниях a , конфигурации соседей у них принципиально различны (см. текст). Пунктиром выделены резонансная пара близко расположенных узлов и компактно расположенная тройка

вставленных друг в друга оболочек, содержащих в общей сложности 51 атом (рис. 7.10).

Квазикристаллы идентифицируют и исследуют в основном по дифракции рентгеновских лучей, так же как и обычные кристаллы. Фурье-образ любой функции координат в идеальном кристалле, напри-

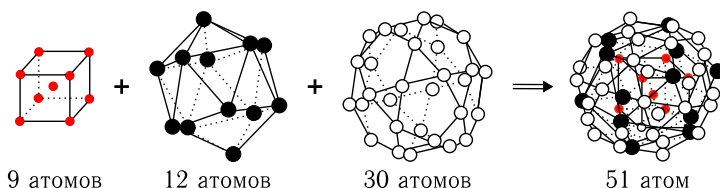


Рис. 7.10. Последовательные оболочки из атомов в псевдо-икосаэдре Макея, являющимся базовым элементом структуры квазикристалла Al–Pd–Mn; суммарное число атомов 51 [10]

мер плотности $\varrho(\mathbf{r})$, представляет собой сумму бесконечного количества узких пиков, в идеале δ -функций:

$$\varrho(\mathbf{r}) = \sum_q \varrho_q \exp(i\mathbf{q}\mathbf{r}). \quad (7.5)$$

Совокупность векторов $\mathbf{q}$ образует решетку в q -пространстве, обладающую той же симметрией, что и исходная решетка атомов. Каждому узлу этой обратной решетки соответствует брэгговский рефлекс на лауэграмме кристалла. Рефлексы тем острее, чем совершеннее кристалл.

Брэгговские рефлексы не являются исключительным свойством кристаллов. В качестве фурье-образа можно изначально взять ряд (7.5), в котором множество векторов $\mathbf{q}$ не обладает трансляционной симметрией, и из него получить обратным фурье-преобразованием функцию $\varrho(\mathbf{r})$, тоже без трансляционной симметрии. Фурье-образами квазикристаллов являются именно такие ряды. При этом ширина брэгговских рефлексов по-прежнему определяется несовершенством структуры: отклонениями локальных конфигураций от идеальной, сбоями дальнего порядка из-за примесей, вакансий и т.п. Чем острее брэгговские рефлексы, тем ближе квазикристалл к идеальному.

В металлических монокристаллах существует корреляция: чем выше качество лауэграммы, тем меньше сопротивление ρ . Эта корреляция отражает волновую природу электронов: улучшение условий распространения рентгеновской волны сопровождается уменьшением рассеяния блоховской волны. В квазикристаллах все происходит ровно наоборот: отжиг, повышающий качество лауэграммы, одновременно увеличивает электросопротивление. Сами значения удельного сопротивления при этом очень велики. Например, в квазикристаллах системы Al–Cu–Ru они достигают при 4 К значений порядка 30 МОм·см. Это примерно в 100 раз больше, чем значение $\rho^* \approx 1/\sigma_{\text{Mott}}$, оцениваемое из концентрации валентных, металлических, электронов n по формулам (1.5) и (1.6) из гл. 1.

Наиболее низкая проводимость наблюдается в системе Al–Pd–Re, в которой удельные сопротивления при 4 К устойчиво получают порядка 200 ÷ 300 МОм·см. Слитки этого сплава получают дуговой плавкой смеси очень чистых Al, Pd и Re в атмосфере чистого аргона. Сплав становится икосаэдрическим квазикристаллом после отжига в вакууме при 980 °С в течение 24 часов. Однако и после этого он остается чувствительным к более низкотемпературному отжигу при 600 °С. Такой отжиг в течение 1–2 часов может удвоить или даже утроить сопротивление при 4 К при неизменном качестве или незначительном улучшении лауэграммы.

Квазикристалл Al₇₀Pd_{22,5}Re_{7,5} отличается от Al₇₀Pd₂₂Mn₈ лишь заменой Mn на изовалентный Re. Поэтому и базисные конфигурации Al₇₀Pd_{22,5}Re_{7,5}, вероятно, аналогичны изображенным на рис. 7.10.

На рис. 7.11 температурные зависимости проводимости для квазикристаллов $\text{Al}_{70}\text{Pd}_{22.5}\text{Re}_{7.5}$ показаны как функции $T^{1/3}$ (четыре нижние кривые) или $T^{1/2}$ (три верхние кривые). Для того чтобы различать образцы, а также различные состояния одного образца, полученные при низкотемпературном отжиге, в качестве параметра можно выбрать значение σ_{10} проводимости при 10 К (масштабы горизонтальных осей выбраны так, что обе шкалы совпадают не только при $T = 0$, но и при этой температуре). Как видно из рис. 7.11, для всех измеренных функций $\sigma(T)$ линейная экстраполяция в выбранных шкалах позволяет определить значение $\sigma(0)$.

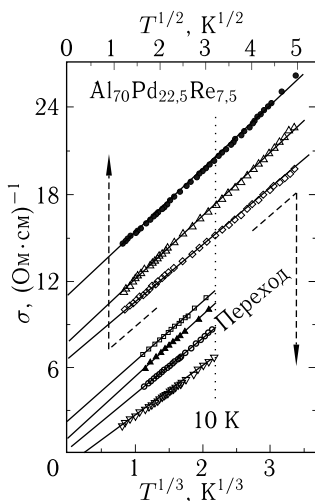


Рис. 7.11. Температурные зависимости проводимости квазикристаллов $\text{Al}_{70}\text{Pd}_{22.5}\text{Re}_{7.5}$, спрямляющиеся при аргументе $T^{1/3}$ в непосредственной близости к переходу металл–изолятор (нижние четыре состояния) и при аргументе $T^{1/2}$ в глубине металлической области (верхние три состояния). Состояния можно маркировать, например, по величине удельной проводимости σ_{10} при температуре 10 К [11]

Общая картина эволюции функции $\sigma(T)$ полностью соответствует тому, что должно происходить в окрестности перехода металл–изолятор (ср. с рис. 6.4 и 6.5 в гл. 6, а также с рис. 8.9 и 8.10 в следующей главе). Для трех верхних состояний с $\sigma(0) \gtrsim 6(\text{Ом} \cdot \text{см})^{-1}$ можно считать выполненным соотношение

$$\Delta\sigma \approx \sigma_{10} - \sigma(0) \lesssim \sigma(0). \quad (7.6)$$

Это позволяет считать для них температурозависящую часть проводимости квантовой поправкой. Именно поэтому зависимость $\sigma(T)$ спрямляется в осях $(T^{1/2}, \sigma)$.

Для четырех нижних состояний с $\sigma_{10} \lesssim 12 \div 14(\text{Ом} \cdot \text{см})^{-1}$ соотношение (7.6) меняется на противоположное. Это означает, что эти состояния находятся в критической окрестности перехода металл–изолятор. Именно поэтому зависимости $\sigma(T)$ для этих состояний построены как функции $T^{1/3}$ и спрямляются при таком построении:

$$\Delta\sigma \equiv \sigma(T) - \sigma(0) \propto T^{1/3}. \quad (7.7)$$

Переход металл–изолятор происходит при $\sigma(10\text{ К}) \equiv \sigma_{10} \approx 9 (\text{Ом} \cdot \text{см})^{-1}$, а у состояний с меньшими значениями σ_{10} низкотемпературный транспорт реализуется при помощи прыжковой проводимости. При углублении в область изолятора температурный интервал, в котором наблюдается прыжковая проводимость, становится достаточно широким. Это иллюстрирует рис. 7.12, из которого видно, что в квазикристаллах $\text{Al}_{70}\text{Pd}_{22,5}\text{Re}_{7,5}$ реализуется закон Мотта

$$\ln \sigma \propto T^{-1/4}. \quad (7.8)$$

Степень $T^{-1/4}$ означает, что плотность состояний электронного спектра в окрестности ферми-уровня имеет постоянное значение, отличное от нуля (см. формулы (4.17)–(4.19) и рис. 4.6 в гл. 4).

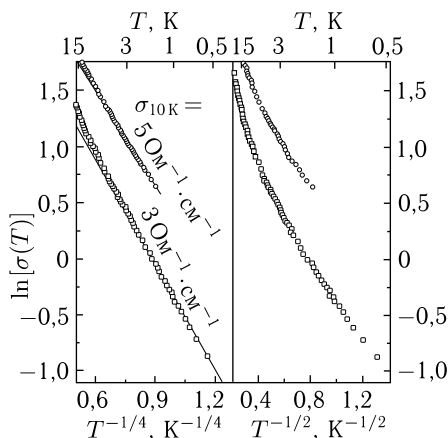


Рис. 7.12. Закон Мотта для проводимости квазикристаллов $\text{Al}_{70}\text{Pd}_{22,5}\text{Re}_{7,5}$, находящихся в области изолятора (около кривых указаны проводимости при 10 К, равные 5 и 3 $(\text{Ом} \cdot \text{см})^{-1}$). Температурные зависимости спрямляются только, когда $\ln \sigma$ построен как функция $T^{-1/4}$ [11]

Таким образом, все происходящее с транспортом в квазикристаллах $\text{Al}_{70}\text{Pd}_{22,5}\text{Re}_{7,5}$ при низкотемпературном отжиге, *улучшающем* условия распространения электромагнитных волновых пакетов, полностью соответствует изложенной в гл. 4–6 картине переходов металл–изолятор в разупорядоченных средах при уменьшении параметра $a_B n^{1/3}$.

Заметьте. Хотя тенденция роста удельного сопротивления при сужении брэгговских рефлексов характерна для многих семейств квазикристаллов, переход металл–изолятор наблюдался пока только в системе Al–Pd–Re. Максимальные значения удельного сопротивления при 4 К в системах Al–Cu–Fe, Al–Cu–Ru и Al–Cu–Mn на один–два порядка меньше, чем в Al–Pd–Re. Следует отметить при этом, что они все равно превышают в 10–100 раз значение ρ^* из формулы (7.1).

Постараемся понять, как же устроен изолятор $\text{Al}_{70}\text{Pd}_{22,5}\text{Re}_{7,5}$. Структура высокоомных квазикристаллов наиболее подробно изучена на примере системы $\text{Al}_{70}\text{Pd}_{22}\text{Mn}_8$, которая отличается от $\text{Al}_{70}\text{Pd}_{22,5}\text{Re}_{7,5}$

лишь заменой Мп на изовалентный Re. Можно считать, что количественные характеристики этих квазикристаллов практически одинаковы.

Итак, основой структуры квазикристаллов $\text{Al}_{70}\text{Pd}_{22}\text{Mn}_8$ (а также $\text{Al}_{70}\text{Pd}_{22.5}\text{Re}_{7.5}$) являются высокосимметричные, близкие к сферической конфигурации из 51 атома (см. рис. 7.10). Согласно дифракционным данным, плотность атомов в этих веществах близка к $6 \cdot 10^{22} \text{ см}^{-3}$. Поскольку атомы переходных элементов забирают часть из трех валентных электронов алюминия, оставшаяся часть «потенциально металлических» электронов составляет чуть меньше двух на атом, т.е. около 10^{23} см^{-3} . Для того чтобы вещество с такой гигантской концентрацией электронов было изолятором, нужно поместить электроны в глубокие потенциальные ямы — ловушки. В интерметаллидных бинарных расплавах ловушками были конфигурации (7.2), в квазикристаллах — высокосимметричные конфигурации рис. 7.10, в которых имеется примерно 90 уровней для бывших валентных электронов. Покинуть ловушку при благоприятных обстоятельствах могут 1–2 электрона с верхних уровней. Поэтому электронная концентрация изначально понижается на 2 порядка, после чего уже можно оперировать более или менее стандартными моделями переходов металл–изолятор.

7.3. Модели электронной структуры в системах с металлическими квазимолекулами-ловушками

Пусть внутри системы делокализованных электронов находится одна атомная конфигурация. Это может быть структурная единица Синтля (7.2), например Cs_4Sn_4 в расплаве с большим избыточным количеством Cs или Sn, или одиночный псевдо-икосаэдр Макея (рис. 7.10) в аморфном сплаве. Электронные уровни ε_i внутри атомной конфигурации, вообще говоря, попадают в зону проводимости сплава-матрицы. Для того чтобы уровень ε_i в зоне отвечал локализованному состоянию, должен существовать потенциальный барьер, не дающий электрону уйти из конфигурации на бесконечность (см. рис. 7.13, а). Если, однако, уровень ε_i расположен достаточно близко к уровню Ферми, т.е. заполнен не полностью, то у электронов имеется конечная вероятность $\Gamma/\hbar$ туннелирования через барьер в зону проводимости и обратно. Поэтому связанное состояние с энергией ε_i называется виртуальным. Время жизни электрона в этом состоянии равно $\hbar/\Gamma$, а ширина уровня — соответственно Γ .

Впервые понятие о виртуальных связанных состояниях на примесях с энергией, находящейся внутри зоны проводимости металла, было введено Фриделем. Затем Андерсон существенно развил эту модель, используя ее для описания локализованных магнитных моментов в металлах. Наконец, модель примесных состояний, находящихся внут-

ри зоны проводимости или валентной зоны, применяется для описания полупроводников типа $A^{IV}B^{VI}$. Таким образом, эта модель хорошо известна и широко используется в физике твердого тела.

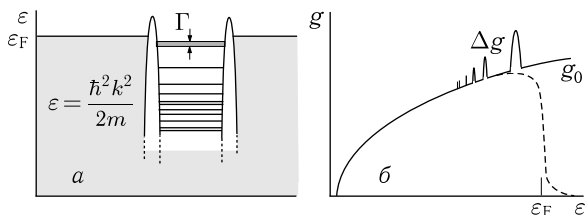


Рис. 7.13. а) Схематическое изображение электронной структуры квазикристалла: уровни базовой конфигурации отделены от ферми-моря барьером. Туннелирование в обе стороны через барьер приводит к уширению верхних уровней конфигурации. б) Плотность состояний $g(\varepsilon)$ электронной системы в квазикристалле. То, что часть электронов находится на квазилокализованных дискретных уровнях внутри конфигураций, уменьшает значения регулярной части g_0 полной плотности состояний g . Пунктиром показано произведение функции $g_0(\varepsilon)$ на фермиевскую функцию распределения

Пусть теперь локальных конфигураций много, но расстояние между ними велико, и они не взаимодействуют между собой. По мере роста концентрации конфигураций $\mathcal{N}$ регулярная часть $g_0(\varepsilon)$ полной плотности состояний $g(\varepsilon)$ уменьшается, а вклад Δg от уровней внутри конфигураций растет. Существует естественный предел концентрации конфигураций. Если n — полная концентрация электронов в сплаве (имеются в виду «валентные» электроны, которые в обычном металлическом сплаве заведомо делокализованы), а N — количество таких электронов в одной конфигурации, то концентрация конфигураций

$$\mathcal{N} \leq \mathcal{N}_{\max} = n/N. \quad (7.9)$$

Величина $\mathcal{N}$ и в сплавах, и в квазикристаллах зависит от химического состава. Например, в расплаве Cs–Sn величина $\mathcal{N}$ увеличивается по мере приближения состава расплава к критическому значению 50:50. Возможно, предел $\mathcal{N}_{\max}$ никогда не достигается: в расплавах из-за высокой температуры должно быть некоторое количество диссоциированных конфигураций; в квазикристаллах между конфигурациями, возможно, существуют аморфные прокладки из атомов сплава. Однако, соотношение

$$\mathcal{N} \approx \mathcal{N}_{\max}, \quad (7.10)$$

судя по экспериментальным данным, может быть реализовано.

Благодаря соотношению (7.10), эффективная концентрация электронов n_{eff} уменьшается в несколько раз. Покинуть ловушку-конфигурацию при благоприятных обстоятельствах могут 1–2 электрона с верх-

них уровней. Остальные как бы уходят в перестроенный ионный остов и присоединяются к электронам внутренних электронных оболочек. Таким образом,

$$n_{\text{eff}} \approx \mathcal{N}_{\text{max}} \approx n/N. \quad (7.11)$$

В квазикристаллах это приводит к уменьшению эффективной концентрации почти на два порядка.

С точки зрения перехода металл–изолятор возможны различные подходы к образовавшейся электронной системе. Можно считать все конфигурации одинаковыми, но хаотически расположенными примесными центрами и рассматривать их в рамках модели структурного беспорядка (см. соответствующий параграф в гл. 5). При этом наличие специфического дальнего порядка в квазикристаллах не убирает беспорядок в расположении ближних соседей. Каждому варианту окружения, в соответствии с моделью структурного беспорядка, описанной в гл. 5, соответствует свой сдвиг верхних электронных уровней данной конфигурации. Поясним это на примере покрытия Пенроуза, для чего обратимся к рис. 7.9. Расстояние от данного узла до соседнего может быть равно стороне ромба a , малой диагонали узкого ромба $a_1 = 0,62a$ и малой диагонали широкого ромба $a_2 = 1,18a$. Но вариантов окружения, определяющего сдвиг уровня конкретного узла, очень много. Например узлы с соседями на наименьшем расстоянии a_1 могут стоять резонансными парами, либо тройками. Узлы 1 и 2 имеют каждый 5 соседей на расстоянии a , но при этом у узла 1 все 5 соседей входят в резонансные пары или в компактные тройки с попарными расстояниями $a_1 < a$, а узел 2 ни одного такого соседа не имеет; узел 3 имеет 6 соседей на расстоянии a , но 3 из них входят в компактную тройку; узел 4 имеет 7 соседей на расстоянии a , но 6 из них составляют две компактные тройки, и т. д. В результате один уровень, изначально одинаковый во всех конфигурациях–узлах, превращается в зону. Являются состояния в этой зоне локализованными или нет, зависит от параметра $a_B n$, в котором величина a_B определяет длину затухания волновой функции вне конфигурационной ямы, а $n = n_{\text{eff}}$.

Специфически низкие значения проводимости, характерные для $\text{Al}_{70}\text{Pd}_{22,5}\text{Re}_{7,5}$, по-видимому обусловлены конкретным сочетанием параметров конфигурационной ямы, которое сделало длину затухания a_B меньшей, чем в других квазикристаллах. При низкотемпературном отжиге $\text{Al}_{70}\text{Pd}_{22,5}\text{Re}_{7,5}$, вероятно, происходит «внутренний ремонт» конфигурационных ям, сопровождающийся уменьшением просачивания волновой функции наружу, т. е. уменьшением длины затухания a_B . Возможно также, что при образовании этого квазикристалла удается ближе подойти к предельной концентрации конфигурационных ям $\mathcal{N}_{\text{max}}$.

Проводимость квазикристаллов можно описывать и на языке гранулированных металлов, рассматривая каждую конфигурацию как грану-

лу, которую в данном случае более уместно называть квантовой точкой в 3D-пространстве. Уход одного электрона с квантовой точки означает появление на ней заряда e , на что требуется энергия порядка

$$\varepsilon_e \approx e^2/r, \quad (7.12)$$

где r — это радиус квантовой точки. Эта величина подобна энергии Хаббарда в теории перехода Мотта. В то же время e^2/r — это кулоновская энергия уединенного металлического шара радиуса r с зарядом e или энергия конденсатора, фигурирующая в теории кулоновской блокады в наноструктурах. На металлической стороне перехода металл-изолятор электрическое поле заряженной точки заэкранировано и энергия (7.12) незначительна. На стороне изолятора нет экранирования свободными носителями и количество заряженных точек определяется сравнением энергии (7.12) с температурой. При $\varepsilon_e \ll T$ количество заряженных точек ν экспоненциально мало:

$$\nu = \mathcal{N} \exp(-\varepsilon_e/T). \quad (7.13)$$

Поскольку проводимость в этих условиях определяется туннелированием между заряженными и незаряженными точками, то ν играет роль количества носителей. Это стандартное рассуждение, используемое при описании гранулированных металлов (см. гл. 8), определяет активационный характер проводимости.

Таким образом два класса конденсированных сред, кратко описанные в данном параграфе, позволяют положительно ответить на вопрос о возможности локализации системы валентных электронов в среде, состоящей только из металлических атомов. Такая возможность реализуется за счет образования молекулообразных конфигураций по крайней мере в двух случаях: в двухкомпонентных расплавах со щелочным металлом в качестве одного из компонентов и в квазикристаллах. В обоих случаях решающим фактором является замена плотности n свободных электронов на $n_{\text{eff}} \ll n$.

7.4. Контрольные вопросы и задачи

1. Проводимость расплава $\text{Cs}_{50}\text{Au}_{50}$ равна 3 (Ом см)^{-1} (см. рис. 7.6), а температурный коэффициент $d \ln \sigma / dT$ при 600°C равен $3,5 \times 10^{-3} \text{ K}^{-1}$. Предполагая, что расплав состоит из комплексов $(\text{CsAu})_N$ (вероятное значение $N = 4$), и что проводимость осуществляется электронами, термически возбужденными из этих комплексов, оценить из данных по проводимости энергию связи электрона в комплексе. (Воспользоваться значением (1.6) и (7.1) для максимального сопротивления ρ^* стандартного металла.)

$$\sigma/\sigma^* = (2N)^{-1/3} \exp(-\varepsilon/3T), \quad \varepsilon \approx 1,6 \text{ эВ.}$$

Список литературы

1. *Hensel F.* // Adv. Phys. **28**, 555 (1979).
2. *Van der Lugt W., Geerstma W.* // Can. Journ. Phys. **65**, 326 (1987).
3. *Calaway W. F., Saboungi M.-L.* // J.Phys. F: Met. Phys. **13**, 1213 (1983).
4. *Meijer J. A., Vinke G. J. B., van der Lugt W.* // J.Phys. F: Met. Phys. **16**, 845 (1986).
5. *Xu R., de Longe T., van der Lugt W.* // Phys. Rev. B **45**, 12788 (1992).
6. *Hoshino H., Schmutzler R. W., Hensel F.* // Phys. Lett. A **51**, 7 (1975).
7. *Schmutzler R. W., Hoshino H., Fischer R., Hensel F.* // Ber. Bunsenges. Phys. Chem. **80**, 107 (1976).
8. Physical Properties of Quasicrystals (ed.: Z. M. Stadnik). Springer, 1999.
9. *Goldman A. I., Kelton R. F.* // Rev. Mod. Phys. **65**, 213 (1993).
10. *Janot C.* // Phys. Rev. B **53**, 181 (1996).
11. *Wang C. R., Lin S. T. et al.* // Journ. Phys. Soc. Japan **67**, 2383 (1998); **68**, 3988 (1999); **69**, 3356 (2000).

Глава 8

ГРАНУЛИРОВАННЫЕ МЕТАЛЛЫ ¹⁾

8.1. Морфология и классификация

Гранулированным будем называть материал, состоящий из случайно расположенных мелких областей (гранул) с существенно различной проводимостью. Случайный потенциал в таком материале обязательно имеет характерные длины, существенно большие межатомных расстояний, вплоть до макроскопических. Рассмотрим пространство с размерностью d и пусть x — доля этого пространства, занятая металлом. Сама по себе величина x еще ни о чем не говорит. Ясно, что проводимость материала с металлическими включениями в виде шариков или в виде тонких нитей совершенно различна при одном и том же x . Морфология материала, под которой мы понимаем здесь форму включений, зависит от множества факторов и чрезвычайно разнообразна. В качестве примера на рис. 8.1 приведены сделанные на сканирующем электронном микроскопе фотографии пленок In, напылявшихся на подложку SiO₂ при комнатной температуре. In не смачивает поверхность, на которую происходит напыление.

Сначала попавшие на подложку атомы, обладающие тепловой энергией, двигаясь вдоль поверхности, собираются в маленькие случайно разбросанные капельки (а). При дальнейшем напылении капельки растут и, соприкасаясь, сливаются в капли большего диаметра (б). Затем металлические области приобретают продолговатую форму. По-видимому при увеличении площади контактов капель с подложкой в их центре возникают участки с сильным сцеплением. При слиянии таких укрупненных капель эти участки играют роль центров пиннинга для перемещающейся массы вещества, понижая симметрию образующихся металлических областей (в). Наконец, на последней стадии перед образованием сплошной пленки, когда относительная площадь $1-x$ зазоров между металлическими областями мала, эти зазоры приобретают форму относительно тонких ветвящихся нитей (г). На это тоже есть свои причины в виде некоторых комбинаций законов смачивания

¹⁾ Обзор транспортных свойств гранулированных металлов можно найти в обзоре [1].

и сцепления напыляемого материала с подложкой, но мы ограничимся констатацией этих морфологических особенностей структуры.

Разобьем d -мерное пространство на элементарные объемы a^d и будем считать, что свойства среды внутри объема не меняются, а свойства двух разных объемов независимы друг от друга. Это означает

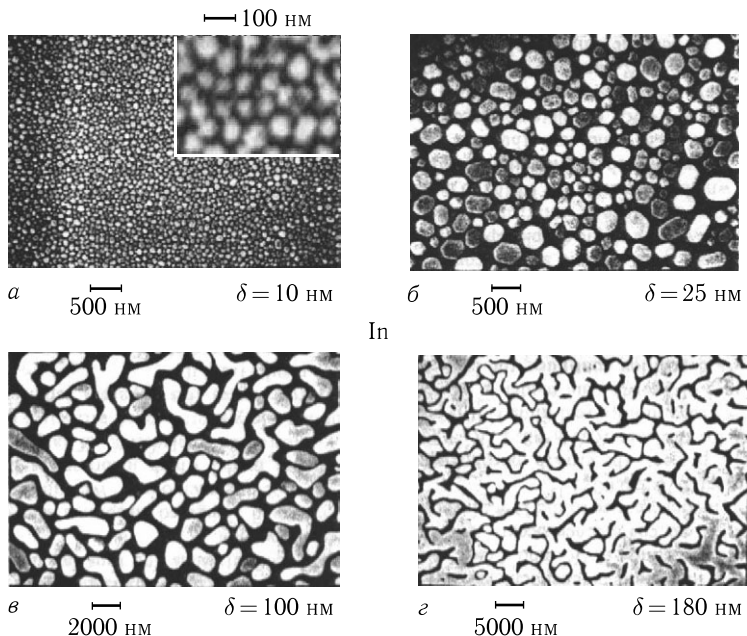


Рис. 8.1. Изображения островковых пленок In, полученные в сканирующем электронном микроскопе [2]. Металл — это светлые области. Под каждой фотографией указаны масштаб и толщина пленки δ . Часть фотографии (а) показана и при большем увеличении

сведение пространственной задачи к задаче на решетке с периодом a и возможность использования простейших моделей теории перколяции. Для структуры на рис. 8.1, а характерный масштаб металлических капель a порядка 0,05 мкм, на рис. 8.1, б он порядка 0,2 мкм. То, что вместе с долей металлического объема x меняется масштаб, мало существенно. Гораздо важнее, что на рис. 8.1, в средний поперечный размер металлических областей меньше, чем их средний продольный размер $b \approx (2 \div 3)a$. Это означает, что на квадратной решетке с периодом порядка a (≈ 1 мкм) появилась корреляция между свойствами b/a соседних узлов.

Математически уменьшение локальной симметрии структуры описывается специфическими корреляторами, введение которых должно сильно усложнить картину, так что простейшие модели теории перколяции — задача связей и задача узлов — становятся непримени-

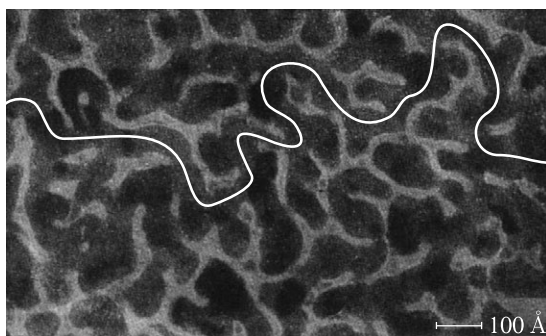
мыми. В этом одно из объяснений того экспериментального факта, что критическое значение $x_c = 0,82 \pm 0,02$ относительной площади покрытия индием поверхности SiO_2 , при котором возникает перколяция, гораздо больше, чем известные критические значения для этих задач. Вторая причина — в потере симметрии между металлическими и неметаллическими областями: если для структур на рис. 8.1, а и 8.1, б можно считать, что области между каплями имеют тот же порядок величины, что и сами капли, то на рис. 8.1, г изолирующие области явно гораздо уже металлических. При этом, однако, они продолжают успешно справляться со своими изолирующими функциями.

Таким образом критическое значение x_c сильно зависит от таких физических факторов, как коэффициент аккомодации падающих на поверхность атомов, величина поверхностного натяжения, величина сил сцепления и т.д. Поэтому при напылении в тех же условиях других металлов получаются другие значения x_c : при напылении Sn получилось $x_c = 0,86 \pm 0,02$, а Pb — $x_c = 0,67 \pm 0,02$.

Заметьте. Наряду с металлическими гранулами в изолирующей матрице можно представить себе и гранулы изолятора в металлической матрице. Но мы, употребляя термин «гранула», будем подразумевать «металлическая гранула». Кроме того, как уже говорилось, гранулированным мы называем и материал со структурой типа рис. 8.1, г, в которой гранулы не являются компактными образованиями.

В системе, представленной на рис. 8.1, роль изолятора, разделяющего металлические гранулы, играет вакуум. Но эту роль может играть и изолятор. Если какие-то металл и изолятор не растворяются друг в друге, то они образуют смесь мелких металлических и изолирующих областей (гранул), называемую кермет. Такая смесь получается, например, при совместном напылении обоих компонентов на изолирующую подложку. Масштаб образующейся структуры контролируется физико-химическими факторами в процессе напыления; в зависимости от них, а также от времени напыления и толщины пленки, могут получаться как двух-, так и трехмерные структуры. На рис. 8.2 представлена электронная фотография кермета $\text{Au} + \text{Al}_2\text{O}_3$ в области существования бесконечного металлического кластера (знак «+» использован для того, чтобы отличать такую гранулированную систему от системы «пленка Au, напыленная на Al_2O_3 »). Здесь также заметна разница в ширинах металлических и изолирующих областей.

Иногда удается сохранить сферическую форму гранул вплоть до большой концентрации металла $x > x_c$. Рис. 8.3, а демонстрирует полученную на просвечивающем электронном микроскопе структуру пленки гранулированного Al в матрице аморфного Ge при концентрации металла $x \approx 0,66$. Видно, что металлический компонент материала состоит из сферических гранул. Специальные измерения позволили



$h = 100 \text{ Å}$

Рис. 8.2. Электронно-микроскопическое изображение гранулированных пленок состава $\text{Au} + \text{Al}_2\text{O}_3$ [1]. Металл — темные области

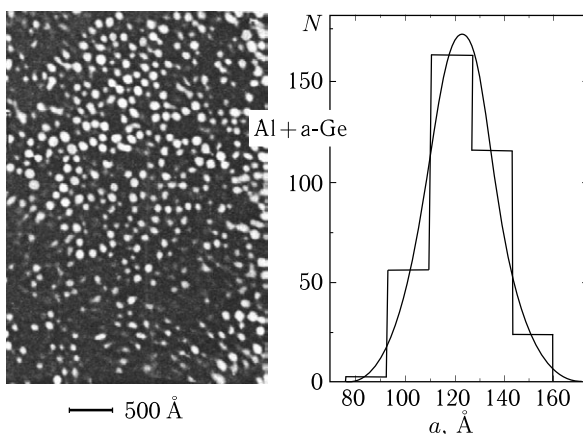


Рис. 8.3. а) Изображения гранулированных пленок $\text{Al} + \text{Ge}$, полученные в просвечивающем электронном микроскопе. Металл — светлые области. б) Гистограмма распределения зерен Al по размерам a и ее аппроксимация нормальным распределением [3]

определить распределение гранул по диаметрам (рис. 8.3, б) — оно оказалось довольно узким.

Во всех упомянутых выше системах на каком-то этапе увеличения относительного объема металла у материала появляется конечная проводимость, т. е. происходит переход изолятор–металл. Такой переход часто называют перколяционным; это название неявно подразумевает, что в основе такого перехода лежат чисто геометрические факторы, так что он является чисто классическим и макроскопическим. Действительно, перколяционные законы инвариантны относительно масштаба, так что можно себе представить перколяцию, например, в системе

металлических шариков от подшипников, случайным образом расположенных на плоскости и зафиксированных застывшим парафином. Но если среди характерных длин в системе есть и достаточно малые, то могут появиться и оказаться определяющими и специфические физические факторы. Мы будем интересоваться именно такими системами.

С другой стороны, если все характерные длины слишком малы, порядка межатомных, то мы возвращаемся к однородно разупорядоченному материалу. Границы между различными классами разупорядоченных систем зависят от того, какими физическими свойствами мы интересуемся.

Поясним это примером, используя важный количественный параметр гранулированной системы: величину расстояния $\delta\varepsilon$ между размерноквантованными уровнями энергии электронов внутри гранул:

$$\delta\varepsilon = (g_F a^3)^{-1} \quad (8.1)$$

(g_F — плотность состояний на ферми-уровне в массивном металле, a^3 — средний объем одной гранулы). Для оценок можно считать, что $\delta\varepsilon \approx 10$ К при $a = 50$ Å.

Если массивный металл — это сверхпроводник с критической температурой T_c и сверхпроводящей щелью Δ , то соотношение

$$\delta\varepsilon \approx \Delta \approx T_c \quad (8.2)$$

определяет минимальный размер изолированной гранулы

$$a_{sc} = (g_F \Delta)^{-1/3},$$

для которой имеет смысл понятие сверхпроводящего состояния. Если $a > a_{sc}$, то сверхпроводящий переход в гранулах происходит при той же температуре, что и в массивном металле, а то, как ведет себя весь материал в целом, зависит от силы взаимодействия между гранулами. Именно это наблюдается в тонких пленках Pb, напыленных на охлажденную до гелиевых температур зеркальную поверхность SiO — см. рис. 8.4, а, взятый из работы [4]. При таком способе напыления атомы свинца собираются в гранулы, достигающие перед слиянием 200 Å в диаметре и 50–80 Å в высоту. У всех пленок толщиной более критической сверхпроводящий переход начинается при одной и той же температуре $T_c \approx 7$ К; у более тонких пленок переход вообще не происходит, но при той же самой температуре $T \approx T_c$ на кривой $\rho(T)$ сохраняется особенность, свидетельствующая о сверхпроводящих переходах в отдельных зернах. При обратном неравенстве, $a < a_{sc}$, гранул, которые могли бы быть сверхпроводящими сами по себе, нет. Материал с точки зрения сверхпроводящего перехода является однородно неупорядоченным, но сверхпроводящее состояние в нем вполне может возникнуть; теперь температура перехода T_c определяется средними характеристиками материала и может плавно меняться вместе с ними. Пленки Pb, напыленные в аналогичных условиях на ту же подложку,

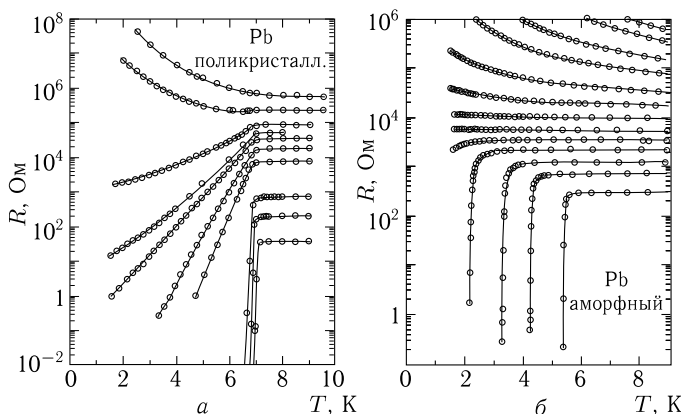


Рис. 8.4. Изменения температурных зависимостей пленок Pb по мере увеличения их толщины (сверху вниз) [4]. а) Пример гранулированной системы: пленки напылены непосредственно на поверхность SiO. Температура сверхпроводящего перехода неизменна. б) Пример мелкодисперсной квазиоднородной системы: пленки напылены поверх промежуточного тонкого слоя аморфного Ge. Кривые $\rho(T)$ демонстрируют корреляцию между сопротивлением в непроводящем состоянии и температурой перехода

но с промежуточным тонким слоем аморфного Ge, демонстрируют корреляцию между температурой T_c и сопротивлением пленки (рис. 8.4, б), т. е. именно такой тип поведения.

Для нормального металла критерий гранулярности иной и зависит от температуры. Соотношение

$$\delta\epsilon \approx T \quad (8.3)$$

определяет минимальный размер гранулы, для которой сохраняет смысл понятие делокализованного электрона. Если в интервал теплового размытия попадает только один электронный уровень, то вообще говоря, правильнее его считать локализованным, а величину a — локализационной длиной.

Умозрительно можно представить себе два типа эволюции гранулированных систем. Первый тип обусловлен изменением величины x . Рисунки 8.1–8.3 иллюстрируют морфологию именно таких систем. Переход металл–изолятор в таких системах имеет как бы перколяционную основу. Но поскольку вместе с x меняется средняя концентрация делокализованных электронов в материале, уместна также ссылка на переход Мотта. Другой тип эволюции выглядит так: при достаточно большом фиксированном x меняются свойства барьеров между гранулами, например, их высота. Количественный параметр такой эволюции можно построить на основе сравнения двух энергий: расщепления $\delta\epsilon$ и интеграла перекрытия волновых функций электронов соседних гранул, который количественно описывает эффективность изолирующих

барьеров. Сравнение с рассуждениями Таулесса, лежащими в основе скейлинговой гипотезы (параграф 6.1), приводит к выводу, что этот тип эволюции имеет много общего с переходом Андерсона в однородно разупорядоченных системах.

На практике произвести такое разделение очень трудно, но условно можно считать, что в следующем параграфе мы будем говорить о гранулированных материалах первого типа, а в последнем — второго типа.

8.2. Кулоновская блокада и переход металл–изолятор

На рис. 8.5, *а* приведены зависимости сопротивления от относительной концентрации металла x в керметах системы $\text{Au} + \text{Al}_2\text{O}_3$ (рис. 8.2), измеренные при двух существенно разных температурах.

Обратите внимание на шкалу на оси ординат: диапазон изменения сопротивления больше 12 порядков.

На графике явно видны две области концентраций x . Область $1 \geq x \gtrsim 0,4$ является металлической: сопротивление ρ сравнительно мало, сравнительно слабо зависит от температуры, увеличиваясь вместе с ней, и постепенно растет с уменьшением x ; где-то вблизи значения $x = x_c \approx 0,38$ находится граница двух областей; наконец, для диэлектрической области $x \lesssim 0,38$ характерен очень резкий рост сопротивления с уменьшением x и очень сильная температурная зависимость $\rho(T)$. Аналогичный график $\rho(x)$ в другой системе, $\text{Ni} + \text{SiO}_2$, приведен на рис. 8.5, *б*. Качественно система ведет себя так же. В частности, и здесь вблизи критического значения x_c производная функции $\rho(T)$ меняет знак (сплошная и пунктирная кривые пересекаются вблизи x_c). Однако само критическое значение x_c другое. О подобном разном значениях x_c на островковых пленках мы уже упоминали.

Стандартное описание в терминах перколяционной модели предполагает, что при концентрациях $x > x_c$ линии тока целиком проходят внутри металлического кластера, а при $x < x_c$ ток должен хотя бы частично проходить через изолятор. Тогда температурную зависимость $\rho(T)$ в области $x < x_c$ должны были бы определять свойства изолятора. Но это верно лишь отчасти.

На рис. 8.6 приведены температурные зависимости материалов гранулированных систем $\text{Au} + \text{Al}_2\text{O}_3$ и $\text{Ni} + \text{SiO}_2$ в изолирующем режиме, т. е. при $x < x_c$. Благодаря тому, что в измеряемом интервале температур сопротивление изменяется на много порядков, удастся определить функциональную зависимость удельного сопротивления $\rho(T)$:

$$\rho(T) = \rho_0 \exp \left(\frac{T_0}{T} \right)^{1/2}, \quad (8.4)$$

надежно отличая ее и от $\exp(T_0/T)$, и от $\exp(T_0/T)^{1/4}$. Функциональной зависимости (8.4) нет у соответствующих массивных изоляторов —

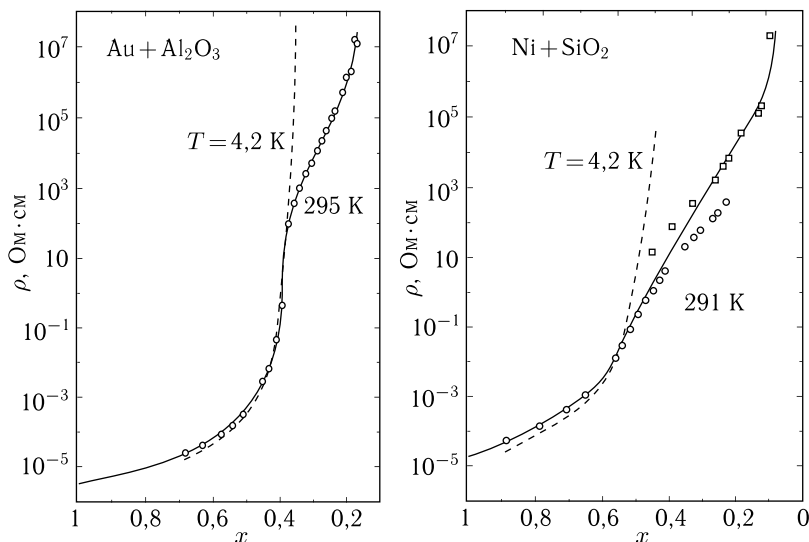


Рис. 8.5. Сопротивление гранулированных пленок $\text{Au} + \text{Al}_2\text{O}_3$ и $\text{Ni} + \text{SiO}_2$ в зависимости от x при двух температурах [1]: пунктирные кривые — при гелиевой, сплошные кривые и значки — при комнатной. На правом графике кружки и квадратики получены в двух разных сериях экспериментов, а сплошная кривая — результат их усреднения.

ни у Al_2O_3 , ни у SiO_2 . К тому же наклон прямых на рис. 8.6, определяемый величиной $T_0^{1/2}$, зависит от величины x . Следовательно, транспорт в обсуждаемых материалах контролируется не только изолятором, но и металлическими гранулами. Это и есть тот экспериментальный факт, который необходимо осмыслить и объяснить.

Теоретическая модель исходит из двух фундаментальных предположений.

Во-первых, между соседними гранулами возможно туннелирование. Здесь теряется инвариантность относительно масштаба решетки, характерная для перколяционных задач, и исключаются системы типа совокупности металлических шариков от подшипников. Не обязательно, чтобы из каждой гранулы было возможно туннелирование во все соседние гранулы. Точнее можно сказать так: совокупность гранул, между которыми возможен обмен носителями посредством туннелирования, должна представлять собой развитый бесконечный кластер. Раз туннелирование — это основной механизм передвижения зарядов в пространстве, то подвижность зарядов пропорциональна вероятности туннелирования:

$$\mu \propto \exp(-\beta a), \quad \beta = \frac{2(2mU)^{1/2}}{\hbar}, \quad (8.5)$$

где a и U — ширина и высота барьера.

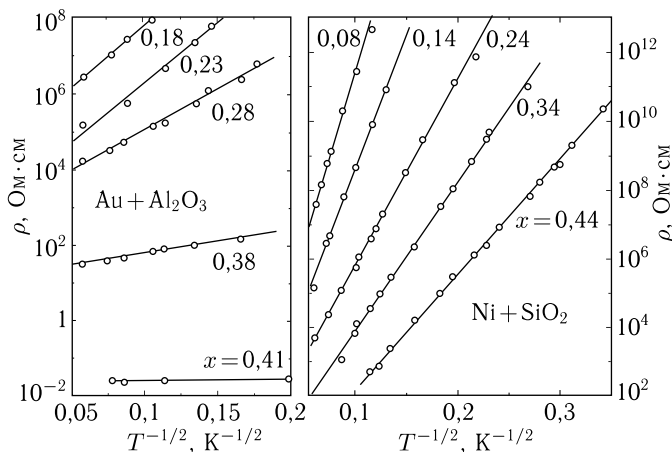


Рис. 8.6. Температурные зависимости сопротивления гранулированных пленок состава $\text{Au} + \text{Al}_2\text{O}_3$ и $\text{Ni} + \text{SiO}_2$ при $x < x_c$ (кроме кривой $x = 0,41$ для материала $\text{Au} + \text{Al}_2\text{O}_3$) [1]

Во-вторых, каждая заряженная металлическая гранула создает электрическое поле в зазоре между собой и соседними гранулами, являясь таким образом обкладкой локального микроконденсатора. Емкость такого конденсатора порядка произведения радиуса гранулы a на диэлектрическую проницаемость κ окружающего его изолятора (емкость уединенного шара):

$$C \approx \kappa a. \quad (8.6)$$

Если заряд в конденсаторе q , то энергия поля в нем $q^2/(2C)$. Поэтому для размещения на грануле одного избыточного электрона, $q = e$, требуется кулоновская энергия $\varepsilon_C \approx e^2/(\kappa a)$. Отсюда следует, что концентрация зарядов n пропорциональна

$$n \propto \exp\left(-\frac{\varepsilon_C}{T}\right) \approx \exp\left(-\frac{e^2}{\kappa a T}\right). \quad (8.7)$$

Заметьте. Материал остается при этом электрически нейтральным, поскольку число электронов и дырок (положительно и отрицательно заряженных гранул) примерно одинаково. Энергия и тех и других отсчитывается от уровня Ферми.

Энергия ε_C отнюдь не мала. Для гранулы размером $50\text{\AA}$ при $\kappa \simeq 10$ она порядка 300 К . Это означает, что при низких температурах туннелирующих носителей экспоненциально мало. Именно это обстоятельство лимитирует проводимость. Отсюда название кулоновская блокада. Оно употребляется чаще применительно к изолированным наноструктурам, таким, как пара туннельных контактов с островком

малой емкости между ними, когда величина ε_C характеризует какую-то конкретную конфигурацию. Однако, неравенство

$$\varepsilon_C \gg T \quad (8.8)$$

может определять и свойства материала как целого.

Формально формулы (8.5) и (8.7) позволяют выделить самые существенные, экспоненциальные множители, входящие в выражение для сопротивления $\rho(T) = \sigma^{-1}(T)$. Поскольку проводимость σ пропорциональна произведению концентрации на подвижность, получаем

$$\rho \propto (n\mu)^{-1} \propto \exp\left(\frac{e^2}{\kappa a T} + \beta a\right). \quad (8.9)$$

Поскольку в показателе экспоненты в выражении (8.9) имеется два слагаемых (параграф 3.3), и длина a в одном из них входит в числитель, а в другом в знаменатель, существует значение

$$a_{\min} = e(\kappa\beta T)^{-1/2}, \quad (8.10)$$

при котором показатель имеет минимум. Значение длины a в реальном материале наверняка имеет дисперсию. Существование минимума означает, что ток в основном будет идти вдоль цепочек из гранул с выделенным значением $a = a_{\min}$, а сопротивление материала будет описываться формулой (8.4) со значением

$$T_0 = 2e\left(\frac{\beta}{\kappa}\right)^{1/2}. \quad (8.11)$$

Прежде, чем обсуждать полученный результат, следует сделать существенную оговорку. Длина a в формулах (8.5) и (8.7) имеет разный смысл: это зазор между гранулами в (8.5) и размер гранул в (8.7). На фотографиях на рис. 8.1–8.3 видно, что они не равны. Но вместо фактически сделанного предположения об их равенстве, можно ограничиться гораздо более реалистичным предположением об их пропорциональности. Это означает, что различные участки после масштабирования становятся статистически одинаковыми. В такой модели основной вывод останется прежним, лишь в выражении (8.11) появится в качестве дополнительного множителя корень из коэффициента пропорциональности. Более того, основной вывод сохранится при любой функциональной связи *при данном x* между размерами гранул a и зазорами a' между ними; лишь бы эти две величины не были статистически независимыми.

Итак, выясняется, что при низких концентрациях металлической фазы, $x < x_c$, ток течет по гранулированному материалу неравномерно, концентрируясь в областях с оптимальным средним размером гранул. Этот оптимальный размер зависит от температуры. Поэтому при изменении температуры линии тока на микроскопическом уровне должны смещаться.

Сравните туннельную проводимость в гранулированной системе с прыжковой проводимостью при наличии кулоновской щели. Одинаковая функциональная зависимость $\ln \rho \propto T^{-1/2}$, одно и то же исходное взаимодействие — кулоновское, схожие механизмы смены с температурой основных токовых путей.

Сходство между этими двумя механизмами не случайно. Если устроить размер гранул к нулю, то они превратятся в примесные центры, которые могут быть либо заряжены, либо электронейтральны. При таком предельном переходе одна задача должна естественно перейти в другую. Но в изоляторе с примесными центрами есть кулоновская щель (параграф 3.3), а в металле с большим количеством примесей из-за кулоновского электрон-электронного взаимодействия появляется минимум плотности состояний на ферми-уровне (параграф 2.4). Чего-то аналогичного следует ожидать и в гранулированном материале. Справедливость этих ожиданий демонстрирует туннельный эксперимент.

На рис. 8.7, *a* представлены туннельные характеристики структуры $\text{Al}-\text{Al}_2\text{O}_3-\text{Ni}+\text{SiO}_2$, в которой одним из берегов туннельного контакта является пленка гранулированного металла, в данном случае состоящая из гранул Ni (металл), не растворяющихся в SiO_2 (изолятор). В эксперименте были использованы пленки толщиной $100 \text{ \AA}$. Поскольку при всех значениях x характерные размеры металлических гранул были меньше $50 \text{ \AA}$, с точки зрения процессов, формирующих электронный спектр, пленка $\text{Ni}+\text{SiO}_2$ представляет собой трехмерную структуру. Плоскость контакта является ее срезом. Металлические гранулы занимают на этой плоскости ту же долю x , что и в объеме. Процесс туннелирования может происходить только в эти, выходящие на плоскость контакта, гранулы.

При больших значениях x , а именно уже при $x = 0,66$ (верхняя кривая), наличие диэлектрических вкраплений несущественно, $\text{Ni}+\text{SiO}_2$ ведет себя как обычный металл, а структура на кривой dJ/dV появляется из-за сверхпроводимости контрэлектрода Al (ср. аналогичный эффект на рис. 2 Приложения Б). Изменения в кривых dJ/dV при меньших x полностью контролируются гранулированным электродом, поскольку с Al ничего не происходит. Абстрагируясь от его неоднородности, можно при помощи стандартной математической процедуры, в соответствии с формулой (Б.1)

$$J(V) \propto \int_{-\infty}^{\infty} g(\varepsilon - eV) g_1(\varepsilon) \left[f\left(\frac{\varepsilon - eV}{T}\right) - f\left(\frac{\varepsilon}{T}\right) \right] d\varepsilon, \quad (8.12)$$

$$f(x) = (\exp x + 1)^{-1},$$

подставляя в качестве $g_1(\varepsilon)$ плотность состояний сверхпроводящего электрода, измеренную при $x > 0,7$, извлечь из каждой эксперимен-

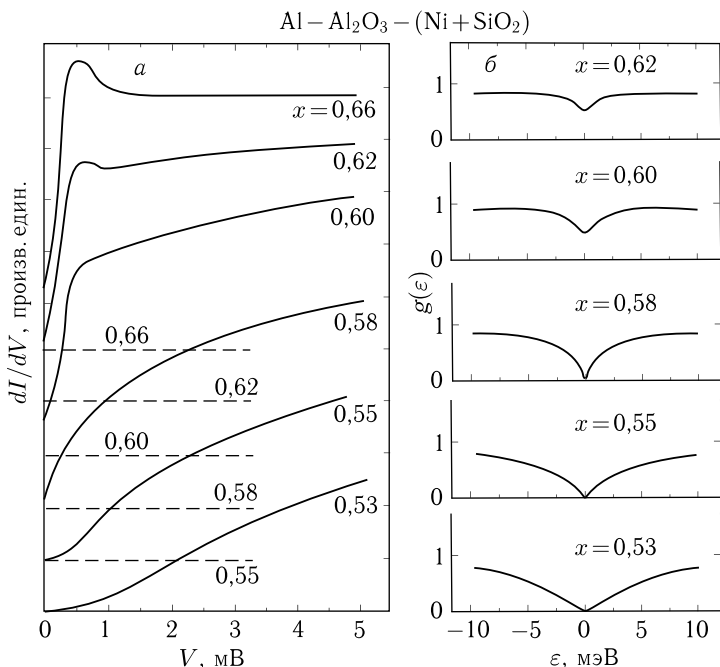


Рис. 8.7. *а*) Туннельные характеристики структуры $\text{Al}-\text{Al}_2\text{O}_3-\text{Ni} + \text{SiO}_2$ с гранулированной пленкой в качестве одного из электродов при различной доли x поверхности (и объема), занимаемой металлическими гранулами, в данном случае гранулами Ni. Кривые смещены для ясности, но для каждой кривой пунктиром показан соответствующий ей ноль, помеченный, как и сами кривые, численным значением x . *б*) Функция $g(\epsilon)$ гранулированной пленки для разных значений x [1]

тальной кривой dJ/dV функцию $g(\epsilon)$ для гранулированного электрода. Результат представлен на рис. 8.7, *б*.

Как видно из сравнения рис. 8.7 с рис. 3 и 4 в Приложении Б, эволюция функции плотности состояний по мере изменения управляющего параметра вблизи перехода металл–изолятор в гранулированном и негранулированном материалах практически одинаковы: в обоих случаях на металлической стороне появляется минимум плотности состояний на ферми-уровне, который превращается в мягкую щель. Согласно рис. 8.7, критическое значение управляющего параметра в системе $\text{Ni} + \text{SiO}_2$ равно $x_c \approx 0,56$. То же значение получается и из кривых рис. 8.5, *б*: именно при этом x меняет знак производная dp/dT .

Таким образом, хотя туннелирование происходит в отдельные гранулы, извлекаемая из эксперимента функция $g(\epsilon)$ отражает состояние всего материала в целом и даже фиксирует происходящий при изменении x переход металл–изолятор. Для объяснения этого обобщим

формулу (8.12) на случай, когда один из электродов неоднородный. То, что туннелирование происходит только в металлическую долю x от полной площади контакта, не очень существенно: это лишь уменьшает эффективную площадь контакта, увеличивая его полное сопротивление в $1/x$ раз. Существенно, что сама металлическая часть среза гранулированного электрода является неоднородной. Часть P поверхности контакта принадлежит бесконечному металлическому кластеру (см. определение мощности бесконечного кластера (А.12) в Приложении А). При туннелировании электрона в эту часть контакта все происходит как в однородном контакте. Эта часть вносит вклад в ток согласно формуле (8.12) с плотностью состояний $g(\varepsilon) = \text{const} = g_0$ (и с учетом уменьшения площади в P раз).

Часть $1-P$ состоит из отдельных гранул размером a с функцией распределения по размерам $D(a)$, нормированной условием

$$\int_0^{\infty} D(a) a^2 da = 1. \quad (8.13)$$

Для того чтобы электрон протуннелировал в такую гранулу, нужна дополнительная энергия $\varepsilon_C \approx e^2/(\kappa a)$. Поэтому эффективная плотность состояний g_a , которую нужно подставить в уравнение (8.12) вместо g , описывая этот процесс туннелирования, равна

$$g_a(\varepsilon) = \begin{cases} 0, & |\varepsilon| < \varepsilon_C(a), \\ g_0, & |\varepsilon| > \varepsilon_C(a) \end{cases} \quad (8.14)$$

(здесь и далее значение $\varepsilon = 0$ соответствует уровню Ферми). Эффективная плотность состояний g_{eff} для всего контакта равна

$$g_{\text{eff}}(\varepsilon) = P g_0 + (x - P) \int_0^{\infty} D(a) a^2 g_a(\varepsilon) da. \quad (8.15)$$

Именно функция $g_{\text{eff}}(\varepsilon)$ извлекается из экспериментальных данных при помощи уравнения (8.12) и именно она изображена на рис. 8.7, б для разных x .

Заметьте. Выражение (8.15) можно было бы преобразовывать дальше, подставив в него выражение (8.14) для $g_a(\varepsilon)$, выражение из (8.7) для $\varepsilon_C(a)$ и т. д. Это не сделано сознательно, поскольку находящееся в начале этой цепочки выражение для емкости (8.6), которое соответствует емкости уединенного шара, является слишком грубым приближением, пригодным лишь для оценок, и должно уточняться при описании конкретных гранулированных структур.

Но уже из неконкретизированного выражения (8.15) видно, что при исчезновении бесконечного кластера, когда происходит перколяционный переход и P обращается в нуль, становится нулем и плотность состояний на ферми-уровне.

Таким образом, ни транспортные измерения в окрестности перехода металл–изолятор, ни туннельные эксперименты не позволяют различить перколяционный переход в гранулированной системе и, например, переход Мотта в однородно разупорядоченной системе. Из самых общих соображений этого следовало ожидать: на переходе расходится корреляционная длина ξ ; когда мы находимся столь близко от перехода, что

$$\xi > a, \quad (8.16)$$

то наличие гранул становится несущественным. Важно, конечно, чтобы a не было слишком большим (как в системе из шариков от подшипников), иначе окрестность перехода станет нереализуемо малой.

8.3. Фрактально-гранулированные металлы

Структура и топология областей, представляющих в гранулированном материале металл или изолятор, могут быть очень сложными, даже отдаленно не напоминающими компактные гранулы. В качестве примера рассмотрим процессы твердотельной аморфизации некоторых метастабильных сплавов.

Существует ряд сплавов, в частности сплавы на основе сурьмы Sb–Zn, Sb–Cd и Sb–Ga, а также сплав Al–Ge, равновесная фаза которых при атмосферном давлении имеет малое координационное число, низкую плотность и ферми-уровень в запрещенной энергетической зоне, а фаза высокого давления — большое координационное число, высокую плотность и ферми-уровень в области разрешенных энергий. Если сплав сначала нагреть под давлением, а затем камеру высокого давления с находящимся внутри нее сплавом охладить до температуры жидкого азота и при этой температуре снять давление и вскрыть камеру, то мы получим сплав в метастабильном состоянии в фазе высокого давления. При азотной температуре эта металлическая метастабильная фаза может сохраняться сколь угодно долго. Установив полученный образец в держатель с прижимными контактами и измерив температурную зависимость сопротивления $R(T)$, можно обнаружить обычный металлический температурный ход сопротивления и сверхпроводящий переход при гелиевых температурах. При нагреве метастабильное состояние должно перейти в стабильное. Если нагрев не очень быстрый, то трансформация происходит в две стадии. Сначала, при некоторой температуре T_1 , тепловой энергии оказывается достаточно для того, чтобы разрушить метастабильную решетку, но ее слишком мало для роста стабильной фазы. Метастабильная фаза переходит в аморфную, у которой координационное число, плотность и проводимость как у стабильной фазы: эта промежуточная фаза является изолятором. Этот процесс и называется аморфизацией. При дальнейшем нагреве произойдет кристаллизация, и аморфный изолятор

превратится в кристаллический изолятор. Мы здесь не будем обсуждать ни эту вторую стадию, ни очень интересные термодинамические подробности, описывающие всю цепочку твердотельных переходов. Это можно найти в специальных обзорах. Здесь речь будет идти о первой стадии процесса, о превращении метастабильного кристаллического металла в метастабильный аморфный изолятор. Именно на этой стадии получается макроскопически неоднородный материал.

Поскольку при аморфизации удельный объем существенно возрастает (в Sb-Zn и Sb-Ga рост достигает 25 %), возникший в метастабильной фазе зародыш аморфной фазы не может изотропно расти из-за повышения локального давления. Однако возможен рост зародыша в виде плоского диска, когда движущийся фронт подобен лезвию, разрезающему метастабильную фазу. Утолщение слоя позади движущейся кромки происходит медленно, поскольку среда вокруг него сжата. Тем не менее напряжение со временем все же растет, и поэтому углы на слое и неоднородно деформированные области являются зародышами ветвления слоев. Таким образом слои изолирующей аморфной фазы распространяются, ветвясь и почти не утолщаясь, и постепенно заполняют все пространство. Многократное ветвление приводит к фрактальным свойствам этих «кактусов», а токовые каналы, вынужденные их огибать, в процессе аморфизации постепенно утоньшаются и становятся все более извилистыми.

Если очень медленно производить нагрев от азотной температуры, то по электросопротивлению можно четко зафиксировать момент, когда сопротивление начнет расти из-за аморфизации. Аморфизацию можно в любой момент прервать, охладив образец. Поэтому переход в аморфное состояние можно производить ступенчато. И в любом из промежуточных состояний можно измерить температурную зависимость сопротивления $R(T)$ и проверить наличие сверхпроводящего перехода.

Основной результат измерений такой. *Постепенной трансформацией можно увеличить сопротивление образца на несколько порядков, практически не изменив температуру сверхпроводящего перехода T_c .* В качестве иллюстрации на рис. 8.8 приведены кривые $R(T)$ для двух сплавов, $\text{Sb}_{57}\text{Cd}_{43}$ и $\text{Sb}_{50}\text{Ga}_{50}$. Все кривые нормированы на сопротивление $R(6\text{ К})$ образца в данном состоянии при температуре $T = 6\text{ К}$. Сопротивление $R(6\text{ К})$ для каждого состояния свое; поэтому около каждой кривой проставлена величина $q = \lg(R/R_0)_{T=6\text{ К}}$, показывающая, во сколько раз изменилось сопротивление образца в результате аморфизации по сравнению с сопротивлением в начальном состоянии без аморфной фазы. Например, значение $q = 6,9$ означает, что сопротивление увеличилось при $T = 6\text{ К}$ почти на семь порядков. Исходные значения удельного сопротивления $1/\sigma$, соответствующие $q = 0$, разнятся от образца к образцу, оставаясь в пределах $10 \div 100\text{ мкОм} \cdot \text{см}$. Это означает, что длина пробега в исходном «однородно металлическом» материале порядка $l \sim 10k_F^{-1}$.

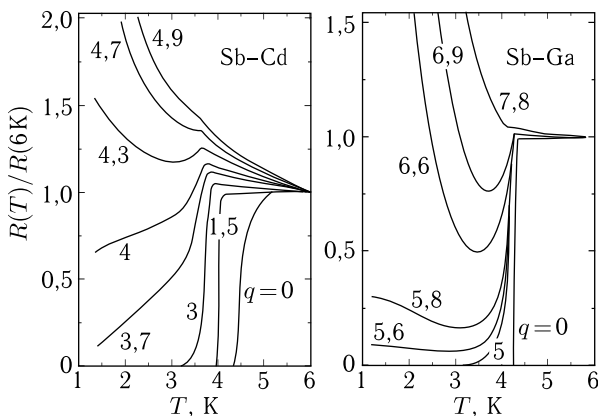


Рис. 8.8. Сверхпроводящий переход в метастабильных сплавах $\text{Sb}_{57}\text{Cd}_{43}$ и $\text{Sb}_{50}\text{Ga}_{50}$ на различных стадиях аморфизации, [6]. Все кривые нормированы на значение сопротивления при $T = 6 \text{ K}$ на соответствующей стадии. Относительное параметра q см. в тексте

Заметьте. В гранулированной системе сверхпроводящий переход в гранулах не обязательно означает падение до нуля общего сопротивления. Все зависит от того, что определяет основное сопротивление вдоль токовых каналов: гранулы или промежутки между ними. Если сопротивление контролируется туннелированием между гранулами, то сверхпроводящий переход в гранулах может даже не уменьшить, а увеличить общее сопротивление. Именно это демонстрируют кривые с большими q на рис. 8.8. С этим же мы встречались и ранее, на рис. 8.4, а.

Согласно соотношению (8.2), постоянство T_c означает, что гранулы остаются большими на всех этапах трансформации. Вместе с тем, они могут становиться рыхлыми, пропитанными проросшим с краев фракталом изолирующей аморфной фазы. Токовый канал внутри такой гранулы в нормальном состоянии подобен траектории частицы, совершающей броуновское движение. Такие материалы мы будем называть фрактально-гранулированными. Чтобы пояснить, откуда берутся при этом большие значения q , обратимся к рис. 8.2, на котором белой линией показан канал, по которому ток может идти слева направо через лабиринтную структуру. Длина канала примерно в $\alpha_L \approx 2$ раза больше длины показанного участка образца, а ширина составляет от ширины этого участка долю порядка $\alpha_h \approx 0,1$. Предположим, что ширина показанного участка — это среднее расстояние между такими сквозными каналами (на языке теории перколяции — это длина корреляции ξ). Тогда, в предположении, что на рис. 8.2 изображена часть двумерного образца, сеть изолирующих прослоек в нем увеличивает его сопротивление в $\gamma = \alpha_L/\alpha_h \approx 20$ раз. Если же на рис. 8.2 показан

срез трехмерного материала, то рост сопротивления в нем порядка $\gamma = \alpha_L/\alpha_h^2 \approx 200$ раз.

Если бы в гранулированном Pb (рис. 8.4) и в сплавах на основе Sb (рис. 8.8) не было сверхпроводящего перехода (или если бы удалось повторить описанные эксперименты в сильном магнитном поле), то в этих системах должен был бы наблюдаться переход металл-изолятор. Вместо этого происходит переход сверхпроводник-изолятор (имеется в виду изменение проводимости при $T \rightarrow 0$ вдоль серий кривых, представленных на рис. 8.4 и 8.8). Этот переход мы здесь обсуждать не будем, поскольку за пределами нашего рассмотрения

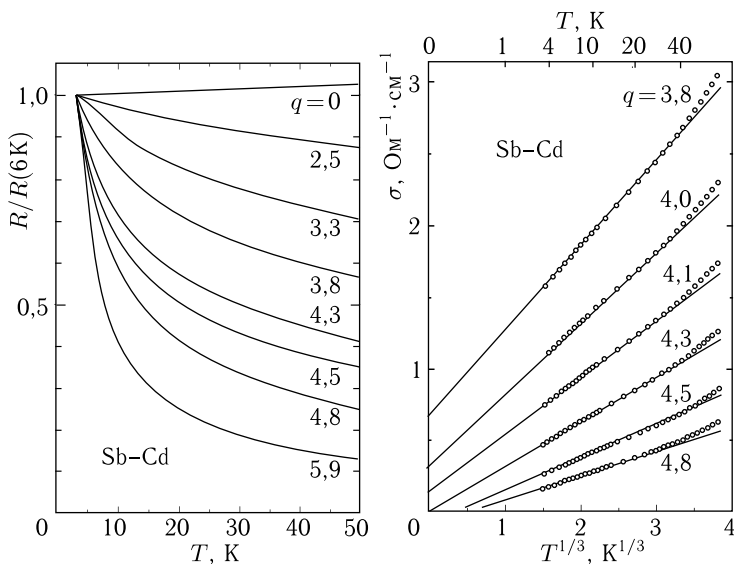


Рис. 8.9. *a*) Температурные зависимости сопротивления метастабильных сплавов $Sb_{57}Cd_{43}$ в нормальном состоянии на различных стадиях аморфизации. Все кривые нормированы на значение сопротивления при $T = 6$ K на соответствующей стадии. Относительно параметра q см. в тексте. *b*) Те же данные в осях ($T^{1/3}$, $\sigma \equiv 1/R$) [7]

осталось само явление сверхпроводимости. Однако можно попытаться увидеть «сквозь него» переход металл-изолятор. Для этого обратимся к схеме окрестности перехода металл-изолятор для однородно разупорядоченных систем (рис. 6.2 из гл. 6).

Воспользоваться этой схемой можно благодаря тому, что критическая область на ней ограничена со стороны высоких температур очень мягким требованием $\tau_{ee} > \tau$: время, за которое происходит расфазировка взаимодействующих электронов, должно быть больше времени упругого рассеяния. Так как $\tau_{ee} \approx \hbar/T$, а $\tau \approx \hbar/\varepsilon_F$, из него следует лишь, что $T \lesssim \varepsilon_F$. Поскольку температура сверхпроводящего перехода

$T_c \ll \varepsilon_F$, то температуры $T > T_c$ с точки зрения перехода металл-изолятор могут считаться не очень высокими. Поэтому для определения положения границы между металлом и изолятором можно, вообще говоря, экстраполировать температурные зависимости проводимости $\sigma(T)$ из области $T > T_c$ к $T = 0$.

На рис. 8.9, *a* приведены кривые $R(T)/R(6\text{ K})$ на разных стадиях аморфизации сплава Sb–Cd. Отображающие изменения сопротивления при $T > 6\text{ K}$, они как бы дополняют кривые $R(T)/R(6\text{ K})$ для того же сплава в сверхпроводящей области $T < 6\text{ K}$ на рис. 8.8, *a*. Эти же данные, перестроенные на рис. 8.9, *б* в осях $(T^{1/3}, \sigma \equiv 1/R)$, выделяют состояние с $(R/R_0)_{T=6\text{ K}} \approx 2 \cdot 10^4$ (т.е. $q = 4,3$) как граничное между металлом и изолятором: линейная экстраполяция экспериментальных данных для этого состояния приводит к прямой, проходящей через начало координат. Следовательно, для этого материала срабатывает та же самая процедура определения точки перехода металл-изолятор, что и в классическом полупроводнике Ge (рис. 4 и 5 в гл. 6) и в квазикристалле (рис. 11 в гл. 7). Тем самым продемонстрирована универсальность этой процедуры и ее пригодность для гранулированного материала.

Вместе с тем, такой переход, по-видимому, может и не наступить. На рис. 8.10 показано аналогичное представление экспериментальных данных в осях $(T^{1/2}, \sigma)$ для самых высокоомных состояний сплава Sb–Ga. Даже при удельных сопротивлениях порядка мегаома экстраполяция к $T = 0$ не обнаруживает перехода в состояние изолятора. Возможно, дело в том, что размеры фрактальных гранул остаются порядка размеров образца, т.е. что не удается удовлетворить неравенству (8.16).

Любопытно, что на температурных зависимостях $\sigma(T)$ для сплава Sb–Ga на рис. 8.10 не видно и тенденции перехода от зависимости $T^{1/2}$ к $T^{1/3}$: даже в самом высокоомном состоянии образца, $q = 8,9$, экспериментальные точки не отклоняются от прямой

$$\frac{1}{R} - \sigma(0) = \beta T^{1/2}.$$

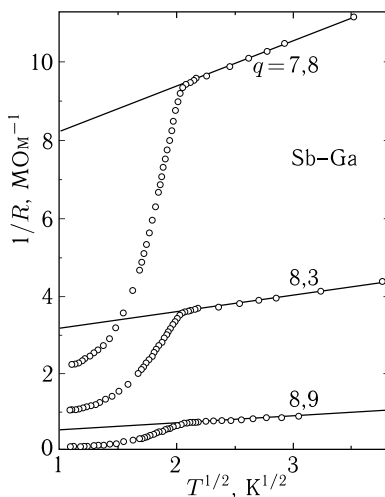


Рис. 8.10. Экстраполяция к $T = 0$ из области $T > T_c$ проводимости метастабильного сплава $\text{Sb}_{50}\text{Ga}_{50}$ на последних стадиях аморфизации дает положительные значения $\sigma(0)$, несмотря на очень малые абсолютные значения σ [6]

По мере уменьшения отсекаемой на оси ординат величины $\sigma(0)$ уменьшается также и наклон прямых β . В отличие от случая, рассмотренного в гл. 7, здесь неравенство (7.6) не ослабевает при понижении температуры, и мы не переходим в область (7.7).

8.4. Контрольные вопросы и задачи

1. Что будет с сопротивлением гранулированного металла, если все линейные размеры системы (размеры гранул b , зазоры между ними d) уменьшить в два раза? Рассмотреть различные случаи, определяемые соотношениями между b и d , их абсолютными значениями по сравнению с межатомным расстоянием, температурой T и т. д.

Список литературы

1. Abeles B., Ping Sheng, Coutts M. D., Arie Y. // Adv. Phys. **24**, 407 (1975).
2. Yu X., Duxbury P. M., Jeffers G., Dubson M. A. // Phys. Rev. B **44**, 13163 (1991).
3. Shapira Y., Deutscher G. // Phys. Rev. B **27**, 4463 (1983).
4. Frydman A. Physica C **391**, 189 (2003).
5. Imry Y., Srongin M. // Phys. Rev. B **24**, 6353 (1981).
6. Гантмахер В. Ф., Зверев В. Н., Теплинский В. М., Цыдынжапов Г. Э., Баркалов О. И. // ЖЭТФ **104**, 3217 (1993).
7. Теплинский В. М., Гантмахер В. Ф., Баркалов О. И. // ЖЭТФ **101**, 1698 (1992).

Глава 9

ЦЕЛОЧИСЛЕННЫЙ КВАНТОВЫЙ ЭФФЕКТ ХОЛЛА ¹⁾

В отсутствие случайного потенциала магнитное поле B делает фи- нитным движение свободного электрона в плоскости, перпендикуляр- ной полю. Соответственно, энергия поперечного движения квантуется (квантование Ландау), и квазинепрерывным остается только спектр движения вдоль поля. Если к тому же электроны находятся в доста- точно тонком слое толщиной b , $b \sim 1/k_F$, перпендикулярном полю, то инфинитное движение вдоль поля тоже невозможно. Тогда электроны должны считаться фактически локализованными: их волновая функция не зависит от времени и отлична от нуля лишь в окрестности класси- ческой электронной орбиты.

В таких условиях привычные представления о локализации непри- менимы. Благодаря магнитному полю, электроны, находящиеся в плос- кой бесконечной яме без какой бы то ни было случайной компоненты потенциала и в отсутствие электрического поля, локализованы на цик- лотронных орбитах. А случайный потенциал, обычно рассматриваемый как возможная причина локализации, приводя в результате рассеяния к переходам с одной циклотронной орбиты на другую, делокализует электроны.

Таков фон предстоящего обсуждения квантового эффекта Холла (КЭХ), одного из самых значительных открытий второй половины XX века в физике твердого тела. Эта тема является существенной ча- стью отдельной обширной области под названием «физика двумерных систем». Ранее мы слегка затронули эту область в гл. 6, обсуждая переходы металл–изолятор. Здесь вся глава посвящена двумерным (2D) системам. Но в соответствии с принципами отбора материала для этой книги мы будем обсуждать только целочисленный кванто- вый эффект Холла, потому что только его можно описать, оставаясь в рамках представлений о невзаимодействующих электронах. Дробный квантовый эффект Холла, которого в этих рамках вообще не должно было бы существовать, останется вне нашего поля зрения. Но и при обсуждении целочисленного КЭХ мы постараемся сосредоточиться на

¹⁾ Подробное обсуждение целочисленного квантового эффекта Холла можно найти в книгах [1, 2]; общие вопросы физики двумерных электронных систем обсуждаются в [3].

тех его аспектах, которые связывают его с основной частью книги: на локализации, переходах металл–изолятор, скейлинге, прыжковой проводимости, перколяции и т. д.

В первом параграфе этой главы мы подробно обсудим те положения классической и квантовой физики, которые необходимы для понимания и описания КЭХ; это и есть тот самый фон, о котором говорилось выше. Затем, после изложения основных экспериментальных фактов, составляющих суть КЭХ, мы построим самосогласованную картину, объясняющую это явление. В последних параграфах речь будет идти о фазовых переходах, сопровождающих КЭХ.

9.1. Спектр и динамика двумерных электронов в сильном магнитном поле ¹⁾

Рассмотрим тонкую бесконечную однородную металлическую пленку в плоскости, перпендикулярной оси Oz . В постоянном магнитном поле $\mathbf{B} \parallel Oz$ траекторией классического движения электронов, не имеющих компоненты скорости вдоль магнитного поля, является окружность. Если в дополнение к магнитному полю, направленному вдоль оси z , на электрон действует электрическое поле $\mathbf{E} \parallel Ox$, то на круговое движение с частотой $\Omega = eB/mc$ накладывается дрейф с дрейфовой скоростью $\bar{v} = c(E/B)$ в направлении $Oy \perp \mathbf{B}, \mathbf{E}$ (здесь c — скорость света, m — эффективная масса). Из-за того, что средняя скорость $\bar{\mathbf{v}} \perp \mathbf{E}$, линейная локальная связь между полем $\mathbf{E}$ и током $\mathbf{j} = ne\bar{\mathbf{v}}$ описывается не скаляром, а тензором (n — концентрация делокализованных носителей). Поэтому при описании линейных транспортных свойств двумерного электронного газа компоненты тока $\mathbf{j}$ и электрического поля $\mathbf{E}$ связывают между собой при помощи тензоров проводимости $\hat{\sigma}$:

$$\begin{aligned} j_x &= \sigma_{xx}E_x + \sigma_{xy}E_y, \\ j_y &= \sigma_{yx}E_x + \sigma_{yy}E_y, \end{aligned} \quad \sigma_{xx} = \sigma_{yy}, \quad \sigma_{yx} = -\sigma_{xy}, \quad (9.1)$$

или сопротивления $\hat{\rho}$:

$$\begin{aligned} E_x &= \rho_{xx}j_x + \rho_{xy}j_y, \\ E_y &= \rho_{yx}j_x + \rho_{yy}j_y, \end{aligned} \quad \rho_{xx} = \rho_{yy}, \quad \rho_{yx} = -\rho_{xy}. \quad (9.2)$$

Эти тензоры являются взаимно обратными, так что для их компонент справедливы соотношения

$$\sigma_{xx} = \frac{\rho_{xx}}{\rho_{xx}^2 + \rho_{xy}^2}, \quad \sigma_{xy} = \frac{\rho_{xy}}{\rho_{xx}^2 + \rho_{xy}^2}. \quad (9.3)$$

¹⁾ В этом параграфе обсуждаются вопросы классической физики металлов, которые можно найти также в книгах [4, 5].

В приближении времени релаксации τ для свободного электронного газа выражения для σ_{xx} и σ_{xy} можно написать непосредственно из уравнения движения свободного электрона, учтя рассеяние введением силы эффективного трения $-m\mathbf{v}/\tau$:

$$m\dot{\mathbf{v}} = \frac{e}{c} [\mathbf{v}\mathbf{B}] + e\mathbf{E} - \frac{m\mathbf{v}}{\tau}. \quad (9.4)$$

Из уравнения (9.4) с учетом соотношений (9.1) и того, что постоянный ток $\mathbf{j} = ne\bar{\mathbf{v}}$, следует

$$\sigma_{xx} = \frac{ne^2\tau}{m} \frac{1}{1 + \Omega^2\tau^2}, \quad \sigma_{xy} = \frac{ne^2\tau}{m} \frac{\Omega\tau}{1 + \Omega^2\tau^2}. \quad (9.5)$$

Из соотношений (9.5) следует, что в *идеальной* 2D-системе без примесей, когда $\tau = \infty$, продольная проводимость бесконечной пленки $\sigma_{xx} = 0$, а холловские проводимость и сопротивление равны соответственно

$$\sigma_{xy} = \frac{ne\bar{v}}{E} = \frac{nec}{B}, \quad \rho_{xy} = \frac{B}{nec}. \quad (9.6)$$

Обычно малая проводимость означает большое сопротивление и наоборот. Между тем, из формул (9.3) следует, что равенство нулю σ_{xx} при $\sigma_{xy} \neq 0$ означает, что и сопротивление $\rho_{xx} = \sigma_{xx} = 0$. Однако проводящая область всегда ограничена в направлении Oy , и это обстоятельство требует тщательного анализа. К тому же в *реальной* 2D-системе всегда есть рассеивающие центры (при конечной температуре роль таких центров могут играть фононы). Поэтому равенство проводимости нулю всегда приближенное: $\sigma_{xx} \approx 0$.

Пусть пленка с рассеивающими центрами имеет вид полосы, вытянутой вдоль оси Ox , и пусть граничные условия задают компоненту поля E_x и ток $\mathbf{j} = (j_x, 0)$. Дрейф вдоль оси Oy приводит к появлению зарядов на краях полосы и к появлению компоненты электрического поля E_y вдоль этой оси. Поэтому электрическое поле в пленке будет иметь обе компоненты: $\mathbf{E} = (E_x, E_y)$. Ток вдоль $\mathbf{E}$ появляется только за счет рассеяния, когда центр круговой орбиты сдвигается вдоль поля. Отношение компонент E_x и E_y электрического поля и величина σ_{xx} определяются частотой рассеяния $1/\tau$.

Край образца существенен благодаря еще одному обстоятельству — вблизи края электроны чаще рассеиваются. В уравнения (9.1) и (9.2) входят плотности тока вдали от краев полосы. Определяя их через геометрические размеры полосы и величину тока, измеряемую снаружи, следует помнить о возможности существования дополнительных краевых токов, не контролируемых частотой $1/\tau$. Рассмотрим классические орбиты, лежащие в плоскости пленки вблизи ее края. Электрон, достигший края и рассеявшийся на нем один раз, обязательно столкнется с ним на следующем витке. В результате таких столкновений электроны совершают инфинитное движение вдоль края, независимо от того, является ли рассеяние на краю зеркальным или диффузным (см. рис. 9.1). Это приводит к тому, что при некоторых специфических

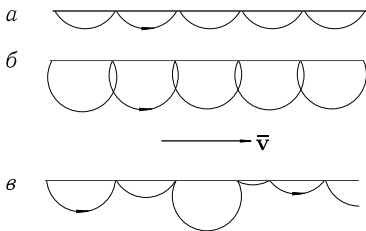


Рис. 9.1. Движение классических электронов вдоль края пленки в нормальном магнитном поле при зеркальном отражении от края (а и б) и при диффузном отражении (в). Во всех случаях средняя скорость электронов $\bar{v}$ отлична от нуля, направлена вдоль края и в первом приближении не зависит от электрического поля

условиях ток вдоль приложенного извне электрического поля в чистом металле течет в основном вдоль его поверхности. На возможность вытеснения постоянного тока к поверхности металла при малой частоте рассеяния в объеме впервые указал Азбель еще в 1963 году. Это явление получило название статического скин-эффекта. Как станет ясно ниже, квантовый аналог статического скин-эффекта играет существенную роль в КЭХ.

К вопросу о терминологии.

Уже на основании классического описания можно сказать, что вдали от краев идеальной пленки электроны локализованы на магнитных круговых орбитах, но что эта локализация мягкая: в сколь угодно малом электрическом поле электроны делокализуются в направлении, перпендикулярном полю. Кроме того, у краев электроны делокализированы в том смысле, что они могут совершать инфинитное движение вдоль края.

Обратимся теперь к квантовому описанию. Идеальный спектр бесспиновых электронов в магнитном поле может быть представлен в виде суммы энергий поперечного $\varepsilon_{\perp}(N)$, зависящего от квантового числа N , и продольного $\varepsilon_{\parallel}(k_z)$ движений:

$$\varepsilon(N, k_z) = \varepsilon_{\perp}(N) + \varepsilon_{\parallel}(k_z) = \hbar\Omega \left(N + \frac{1}{2} \right) + \frac{\hbar^2 k_z^2}{2m}, \quad N = 0, 1, 2, \dots, \quad (9.7)$$

где k_z — волновой вектор вдоль направления магнитного поля $\mathbf{B}$. Кратность вырождения уровней с данными N и k_z есть

$$\gamma = \frac{S}{2\pi r_B^2}, \quad (9.8)$$

где S — площадь перпендикулярного полю сечения области, занятой электронами, а r_B — магнитная длина,

$$r_B = \left(\frac{\hbar c}{e|B|} \right)^{1/2}. \quad (9.9)$$

Если область, занятая электронами, имеет форму пленки толщиной b , лежащей в плоскости (x, y) , а поле $\mathbf{B}$ направлено вдоль оси z , то величина k_z в уравнении (9.7) тоже квантуется: $k_z = (2\pi/b)N_z$ ($N_z = 1, 2, \dots$) и $\varepsilon(N, k_z) \rightarrow \varepsilon(N, N_z)$. Квазинепрерывный одномерный спектр $\varepsilon_{\parallel}(k_z)$ в каждой из магнитных подзон с магнитным

квантовым числом N превращается в набор дискретных неэквидистантных уровней:

$$\varepsilon_{\parallel}(k_z) \rightarrow \varepsilon_{\parallel}(N_z) = \frac{\hbar^2}{2m} \left(\frac{2\pi N_z}{b} \right)^2, \quad (9.10)$$

$$\Delta\varepsilon_{\parallel}(N_z) = \varepsilon_{\parallel}(N_z + 1) - \varepsilon_{\parallel}(N_z) = 2\pi^2(2N_z + 1) \frac{\hbar^2}{mb^2}.$$

Спектр становится полностью дискретным, каждый уровень в нем характеризуется двумя квантовыми числами: N и N_z . Вырожденность γ всех уровней одинакова; согласно (9.8), она зависит только от площади S , занимаемой двумерным электронным газом, и от величины магнитного поля B . Плотность состояний такого идеального 2D-газа, нормированная на единицу площади, имеет вид суммы δ -функций:

$$\sum_{N, N_z} \frac{1}{2\pi r_B^2} \delta(\varepsilon - \varepsilon(N, N_z)). \quad (9.11)$$

Коэффициент $n_L = (2\pi r_B^2)^{-1}$, входящий во все члены суммы (9.11), является плотностью электронов на любом полностью заполненном уровне Ландау. Ее можно выразить через квант магнитного потока Φ_0 :

$$n_L = \frac{\gamma}{S} = (2\pi r_B^2)^{-1} = \frac{B}{\Phi_0}, \quad \Phi_0 = \frac{2\pi\hbar c}{e}. \quad (9.12)$$

Равенство (9.12) означает, что на полностью заполненном уровне Ландау на каждый электрон приходится ровно один квант магнитного потока.

Пусть n_3 — концентрация электронов в трехмерном пространстве; соответственно, концентрация электронов, отнесенная к единице площади пленки, равна $n_2 = n_3 b$. Тогда из выражения (9.8) для γ следует, что при температуре $T = 0$ количество занятых уровней определяется числом заполнения ν , равным

$$\nu = \frac{n_2}{n_L} = \frac{2\pi\hbar c n_2}{eB}. \quad (9.13)$$

Целая часть числа ν — это количество полностью заполненных уровней, наличие дробной части указывает на существование частично заполненного уровня и определяет, в какой мере он заполнен. При увеличении поля вырожденность γ растет, и число заполненных уровней уменьшается. Опустошение следующего уровня начинается, когда дробная часть ν обращается в нуль — в полях

$$B_i = \frac{2\pi\hbar c}{e} n_2 \frac{1}{i} = \Phi_0 n_2 \frac{1}{i}, \quad (i = 1, 2, 3, \dots). \quad (9.14)$$

Нас здесь интересуют очень тонкие пленки в сильных полях, когда заполнено всего несколько дискретных уровней. Для обсуждения удобно воспользоваться приближенной картинкой, которую можно назвать «переводом квантового описания на полуклассический

язык». На рис. 9.2, *a* изображено сечение Ферми-сферы $\varepsilon(k_z, k_\perp) = \varepsilon_{F0}$ в k -пространстве в трехмерном случае без магнитного поля. Ферми-радиус в нулевом магнитном поле ε_{F0} , зависящий только от концентрации n_3 , задает на этой диаграмме масштаб энергии. Магнитное поле квантует поперечный импульс $k_\perp$. Целые числа N отмечают его разрешенные значения, а проходящие через них вертикальные линии — это сечения так называемых цилиндров Ландау, используемых в физике металлов при анализе эффекта де Гааза–ван Альфена. Конечная толщина пленки b квантует продольный импульс k_z , его разрешенные значения отмечены числами N_z . Черные точки внутри окружности радиуса ε_{F0} — это заполненные уровни с квантовыми числами (N, N_z) . На рис. 9.2, *a* их девять и еще один уровень $(N, N_z) = (3, 1)$, по всей видимости, заполнен частично (серая точка). Поскольку Ферми-уровень зависит от магнитного поля, осциллируя в окрестности ε_{F0} , т. е. $\varepsilon_F(B) \neq \varepsilon_{F0}$, точно степень заполнения промежуточного уровня по рисунку сказать нельзя; она определяется соотношением (9.13).

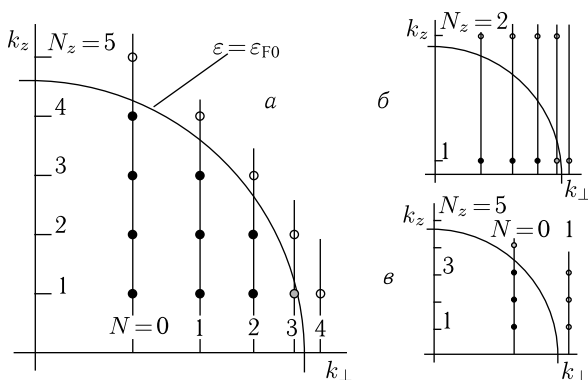


Рис. 9.2. Разрешенные состояния на плоскости $(k_\perp, k_z)$; их энергия пропорциональна квадрату расстояния от начала координат. Черные кружки — полностью заполненные уровни, серый кружок — уровень, заполненный частично, светлые кружки — пустые уровни. *a*) Заполнены уровни с разными, и магнитными N , и размерными N_z , квантовыми числами; *б*) заполнены только уровни с минимальным размерным квантовым числом $N_z = 1$; *в*) заполнены только уровни с минимальным магнитным квантовым числом $N = 0$

При повышении магнитного поля происходит переход к ультраквантовому пределу, в котором из дискретных квантовых уровней заполнен только один уровень $(0, 1)$. Как видно из рис. 9.2, этот переход, в зависимости от величины b , может происходить по-разному. В ультратонких пленках и гетеропереходах, где $\Delta\varepsilon_{||}(1) \gtrsim \varepsilon_{F0}$, заполнены только уровни $(N, 1)$ в горизонтальном ряду — см. рис. 9.2, б. Эксперименты с квантовым эффектом Холла обычно производятся именно на таких объектах. Но возможен и другой случай, когда несколько уровней $\varepsilon_{||}(N_z) < \varepsilon_{F0}$,

например, когда $\varepsilon_{\parallel}(N_z) < \varepsilon_{F0}$ при $N_z < 4$ — рис. 9.2, *а, в*. Тогда при повышении поля на каком-то этапе останутся заполнены только уровни типа $(0, N_z)$ вертикального ряда. Поскольку вырожденность уровней не зависит от квантовых чисел N и N_z , формулы (9.8), (9.13) и (9.14) применимы в обоих случаях. К обоим случаям применимо также все дальнейшее содержание этого параграфа.

Наличие спина удваивает количество дискретных уровней. Наборы, изображенные на рис. 9.2, реализуются для обеих проекций спина, но они сдвинуты друг относительно друга на величину зеемановского расщепления. Дополнительное увеличение числа уровней и усложнение спектра возможно также из-за многодолинности электронного спектра в кристалле.

Получившемуся набору дискретных уровней соответствуют классические круговые орбиты радиуса $r_N \approx \sqrt{N} r_B$ с фиксированными центрами. (Если $N_z \neq 1$, то вращение по орбите сопровождается пе-

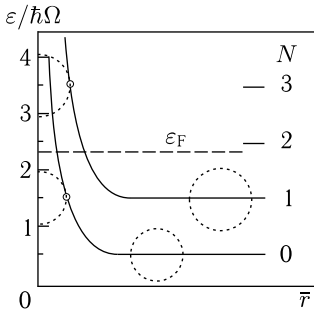


Рис. 9.3. Энергия 2D-электронов на дискретных уровнях Ландау в зависимости от расстояния $\tilde{r}$ от стенки прямоугольной потенциальной ямы до самой удаленной точки классической орбиты. Пунктиром изображены квазиклассические образы волновых функций — полные окружности вдали от края и части окружности вблизи края. Белыми точками отмечены состояния $\tilde{r} = r_N$, в которых они представляют собой точно полуокружности

риодическим движением поперек пленки, между ее поверхностями, но остается финитным.) Приложенное к пленке постоянное однородное электрическое поле $\mathbf{E} \parallel Oy$ превращает дискретные уровни в полосы: $\varepsilon(N, N_z) \rightarrow \varepsilon(N, N_z) + Ey$. Соответственно, как мы видели из уравнений (9.1)–(9.6), классическое движение становится инфинитным в направлении поперек $\mathbf{E}$.

Введение конечной площади ямы S в формуле (9.8) для вырожденности γ учитывает ограниченность проводящей 2D-области. Будем для простоты полагать, что яма прямоугольная, т. е. что стенки у нее вертикальные. Сама по себе вырожденность γ обусловлена трансляционной инвариантностью, т. е. равноценностью орбит с разными координатами центра. Когда расстояние от центра орбиты до края меньше классического радиуса орбиты r , наличие стенки ограничивает волновую функцию, и энергия электронного уровня повышается (см. рис. 9.3, где по оси абсцисс отложено расстояние $\tilde{r}$ от края до самой удаленной точки классической орбиты; предполагается, что $N_z = 1$; это предположение соответствует случаю (б) на рис. 9.2). В частности, энергия орбиты на уровне N с центром точно на стенке, классическим аналогом которой

является полуокружность, равна энергии орбиты вдали от стенки на уровне $N' = 2N + 1$,

$$\varepsilon(N, \bar{r} = r_N) = \varepsilon(N', \bar{r} \gg r_B).$$

Главный эффект увеличения энергии уровней вблизи стенки в том, что у стенки все уровни Ландау пересекают Ферми-уровень и поэтому могут вносить вклад в проводимость за счет краевого тока. Превращение дискретных уровней в полосы у краев пленки немедленно приводит к появлению инфинитных орбит (рис. 9.1).

Описанная идеальная квантовая картина заведомо нарушается даже в самой чистой системе наличием случайного потенциала $U(\mathbf{r})$. Поскольку магнитное поле задает характерный масштаб длины (9.9), можно говорить о двух предельных типах случайного потенциала: потенциал с крупномасштабными флуктуациями с характерными размерами $\zeta \gg r_B$ и короткодействующий потенциал, у которого $\zeta \ll r_B$. И тот, и другой снимают вырождение уровней (N, N_z) и приводят к их уширению. Вопрос в том, каковы волновые функции в этих миниполосах, в которые превращаются уровни Ландау.

Рассмотрим более подробно крупномасштабные флуктуации потенциала $U(\mathbf{r})$. Они приводят к тому, что на электрон действует сила ∇U , лежащая в 2D-плоскости и однородная на размерах классической орбиты. Воспользуемся уравнением (9.4), заменив $\mathbf{E}$ на ∇U и положив $\tau = \infty$. Под влиянием силы ∇U электрон дрейфует в направлении, перпендикулярном ∇U , т. е. все время движется вдоль эквипотенциальных линий функции $U(\mathbf{r})$. Следовательно, каждый электрон находится все время в своем постоянном потенциале U , и именно от него следует отсчитывать энергии уровней (N, N_z) . Поэтому возникает «неоднородное уширение» исходных δ -образных вырожденных уровней (N, N_z) .

Эквипотенциальные линии функции $U(\mathbf{r})$ вблизи локальных экстремумов представляют собой замкнутые контуры, охватывающие эти экстремумы. Направление движения электронов по таким контурам зависит от того, охватывают ли они локальные максимумы U_+ или минимумы U_- . Характерные размеры контуров $\tilde{R}$ являются расстояниями, на которых локализованы волновые функции соответствующих электронов.

К вопросу о терминологии: Волновую функцию электрона в идеальном 2D-слое в сильном нормальном магнитном поле мы назовем условно или мягко локализованной, потому что, как мы видели, сколь угодно малое электрическое поле делокализует электрон в направлении, перпендикулярном полю. В отличие от этого, квазиклассические состояния на эквипотенциальных линиях конечных размеров $\tilde{R}$ истинно или жестко локализованы: электрическое поле лишь деформирует эквипотенциальную линию, причем тем меньше, чем слабее поле. Таким образом крупномасштабные флук-

туации потенциала превращают мягкую локализацию в жесткую по крайней мере для части электронов на уровне Ландау.

В хаотическом 2D-потенциале имеется перколяционный порог, на котором бесконечного кластера не существует, а линии эквипотенциалей превращаются в бесконечную сетку с бесконечным числом самопересечений (см. Приложение А). В седловых точках квазиклассическая скорость электрона равна нулю; поэтому формально можно считать электрон на участках эквипотенциальной линии между точками ее самопересечения мягко локализованным. Учитывая, что вырождение в точках самопересечения снимается электрическим полем и что преодоление седловых точек возможно также за счет туннелирования, часто электроны на перколяционном уровне сразу называют делокализованными. Таким образом, в каждой уширенной миниполосе Ландау имеется один перколяционный уровень, на котором электроны делокализованы. При статистически симметричном случайном потенциале он расположен в центре миниполосы.

Заметьте. Вообще говоря, один уровень не означает одно состояние, потому что уровень может быть вырожденным. В контексте КЭХ этот вопрос очень не простой; мы будем ниже к нему возвращаться (см. формулу (9.28) и дискуссии в тексте в связи с рис. 9.21).

Фактическая ширина $\delta\epsilon$ интервала энергий с делокализованными волновыми функциями зависит от более тонких процессов, например, от туннелирования между двумя квазиклассическими траекториями, близкой подходящими друг к другу в окрестности седловой точки (магнитный пробой). По сути дела величина $\delta\epsilon$ является неопределенностью энергии любого из делокализованных состояний.

Истинно локализованные состояния возникают и в короткодействующем случайном потенциале. Рассмотрим короткодействующий потенциал $\zeta \ll r_B$, создаваемый изолированными примесями. Энергия электрона, двигающегося в непосредственной окрестности примеси, меняется, что приводит к уширению уровня Ландау. Но этим влияние примеси не ограничивается. Разобраться в классификации уровней, возникающих в окрестности примеси, помогает модель водородоподобной примеси в трехмерном пространстве. В отсутствие магнитного поля примесь образует под континуумом делокализованных состояний дискретный набор локализованных состояний, которые описываются набором квантовых чисел, как в атоме водорода. Если наложить сильное магнитное поле, то множество делокализованных состояний распадется на набор одномерных магнитных зон Ландау, а локализованные состояния окажутся под каждой из них: ведь, оторвавшись от примеси, электрон может оказаться в любой из подзон Ландау. Уменьшение толщины пленки и переход к двумерной системе преобразует подзоны

Ландау в дискретные уровни, на которых электроны условно локализованы, но изначальные истинно локализованные состояния сохраняются. Поэтому примеси не только уширяют изначальные дискретные уровни (N, N_z), но и создают около них дискретные (и соответственно, локализованные) уровни.

9.2. Экспериментальные наблюдения целочисленного КЭХ

Для транспортных измерений наиболее распространена измерительная конфигурация, которая называется холловский мостик (см. рис. 9.4, а) и в которой ток J_{12} пропускается через торцевые контакты 1 и 2 (вдоль оси x), а между остальными контактами измеряются разности потенциалов. С ее помощью измеряются компоненты тензора $\hat{\rho}$: ρ_{xx} по разности потенциалов на контактах, расположенных вдоль одного края, и ρ_{xy} при помощи контактов, расположенных напротив друг друга. Формула для ρ_{xx} , написанная на рис. 9.4, а, верна в предположении, что ток в мостике распределен равномерно. Поскольку в условиях КЭХ такой уверенности, вообще говоря, нет, правильное считать, что при КЭХ холловский мостик измеряет $R_{xx} = V_{34}/J_{12}$.

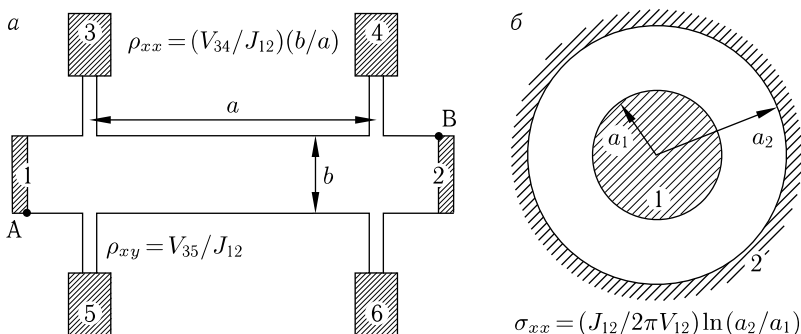


Рис. 9.4. а) Холловский мостик, позволяющий измерять компоненту ρ_{xx} тензора сопротивления по разности потенциалов между контактами 3 и 4 или 5 и 6 и компоненту ρ_{xy} по разности потенциалов между контактами 3 и 5 или 4 и 6 при токе J_{12} через торцевые контакты 1 и 2; относительно точек А и В см. ниже в тексте. б) Диск Корбино с двумя кольцевыми электродами радиусов a_1 и a_2 , позволяющий измерять компоненту σ_{xx} тензора проводимости по разности потенциалов V_{12} между коаксиальными электродами и току J_{12} между ними

Кольцевая структура под названием «диск Корбино», изображенная на рис. 9.4, б, позволяет измерять компоненту σ_{xx} тензора $\hat{\sigma}$.

Двумерный электронный газ обычно создают в одном из двух типов кристаллических структур. Во-первых, это может быть поверх-

ность полупроводника, вблизи которой при помощи перпендикулярного электрического поля создают изгиб зон. Такое устройство называется полевой или МОП (металл-окисел-полупроводник)-структурой. Меняя напряжение на металлическом электроде (затворе), можно управлять концентрацией двумерного газа. Двумерный электронный газ можно создать также внутри кристалла, на границе двух специальных образом подобранных полупроводников. Изгиб зон вблизи границы такой гетероструктуры возникает при выравнивании ферми-уровней с двух сторон от границы. Концентрацию электронов в гетероструктуре обычно тоже можно менять при помощи дополнительного затвора, подавая на него потенциал V_g . Поэтому менять число заполненных уровней (9.13) можно не только при помощи магнитного поля B , но и при помощи потенциала V_g на затворе относительно 2D-газа. Если затвор находится на расстоянии s от плоскости с 2D-электронами, то изменение напряжения на затворе ΔV_g и изменение электронной концентрации Δn связаны линейной зависимостью

$$\Delta V_g = 4\pi \Delta n e s. \quad (9.15)$$

На рис. 9.5 приведены результаты эксперимента [6], по сути дела открывшего КЭХ и положившего начало его изучению. На кривых видны три главные особенности, составляющие суть этого явления: наличие плато на кривых $\rho_{xy}(V_g)$ в окрестности значений V_g , при

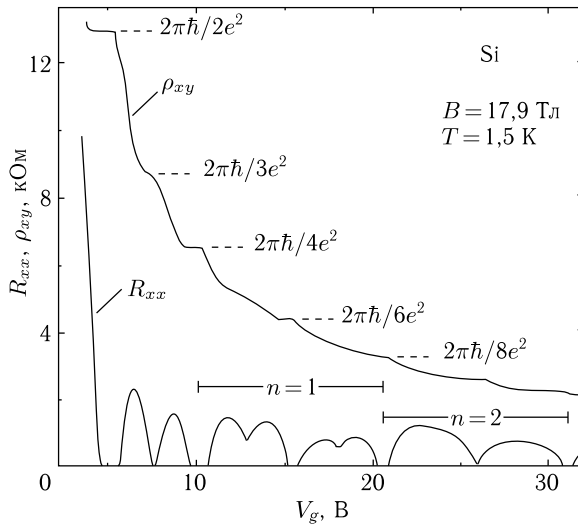


Рис. 9.5. Магнетосопротивление R_{xx} и холловское сопротивление ρ_{xy} в зависимости от напряжения на затворе V_g в полевой структуре на поверхности Si [6]. Температура $T = 1,5$ К, магнитное поле 17,9 Тл. Размеры устройства в обозначениях рис. 9.4, $a = 130$ мкм, $b = 50$ мкм, ток $J_{12} = 1$ мкА

которых магнитное поле B удовлетворяет соотношению (9.14) для полей B_i ; квантованные значения ρ_{xy}

$$\rho_{xy} = \frac{1}{i} \left(\frac{2\pi\hbar}{e^2} \right) \quad (i = 1, 2, 3, \dots) \quad (9.16)$$

на этих плато; обращение $R_{xx}(V_g)$ в нуль в окрестности этих же значений V_g . Эти особенности хорошо видны на рис. 9.6, где окрестность одного из плато показана в увеличенном масштабе (разные величины V_g на рис. 9.5 и на рис. 9.6 объясняются тем, что эксперименты сделаны на разных образцах).

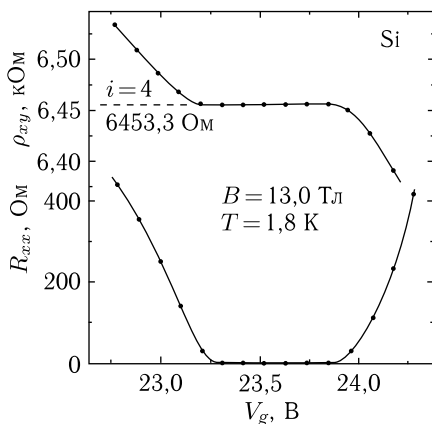


Рис. 9.6. Магнетосопротивление R_{xx} и холловское сопротивление ρ_{xy} в зависимости от напряжения на затворе V_g в полевой структуре на поверхности Si [6]. Температура $T = 1,8$ К, магнитное поле 13 Тл. Размеры $a = 130$ мкм и $b = 50$ мкм те же, что на рис. 9.5, но расстояние s до затвора другое. Точность измерения холловского сопротивления 0,1 Ом

Рис. 9.5 позволяет сформулировать первые вопросы, на которые эксперимент должен ответить для начала серьезной дискуссии о природе КЭХ. Эти вопросы и ответы на них таковы.

(i) *Точность*, с которой выполняется соотношение (9.16). Уже в первой работе [6] было показано, что измеренные значения ρ_{xy} на плато совпадают со значениями (9.16) с точностью не хуже, чем 10^{-4} . Сейчас уже говорят о воспроизводимости получаемых значений ρ_{xy} на уровне 10^{-8} . Это означало возможность применения этого эффекта в метрологии и потребовало выяснения в специальных экспериментах, влияют ли на измеряемые значения ρ_{xy} знак носителей, многодолинность спектра и другие факторы, определяемые кристаллической структурой (оказалось, что не влияют), изучения влияния примесей, температуры, геометрических размеров образцов и величины измерительного тока, влияния электронной концентрации (она тоже оказалась несущественной в довольно широких пределах), исследования изменений ρ_{xy} вдоль пла-

то и т. п. Результаты этих исследований можно найти в специальных обзорах (см. [7, 8]). Нам для дальнейшего достаточно утверждения, что на фоне плавно меняющегося холловского сопротивления $\rho_{xy}(B)$ или $\rho_{xy}(V_g)$ выделяются квантованные значения (9.16).

(ii) *Относительная протяженность плато*, их ширина по сравнению с расстоянием между ними по полю B или по напряжению на затворе V_g . Эта характеристика не обладает устойчивостью квантованных значений холловского сопротивления. Ширина плато зависит и от материала, на основе которого создан 2D-слой носителей, и от примесей, и от температуры. На рис. 9.7 приведены кривые КЭХ на гетероструктуре $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$. Плато занимают большую часть кривой $\rho_{xy}(B)$, так что эта кривая приобретает вид лестницы с горизонтальными ступенями. Влияние температуры на ширину плато демонстрирует рис. 9.8. Соответствующую структуру имеет и функция $R_{xx}(B)$: она отлична от нуля только на интервалах между плато, где имеет вид более или менее узких пиков.

(iii) *Температурная зависимость* продольной проводимости в области плато. Согласно формуле (9.3), в области плато, где $\rho_{xx} \approx 0$, следует также ожидать $\sigma_{xx} \approx 0$. Но эта формула справедлива лишь при пространственно однородном распределении тока и потому одно-

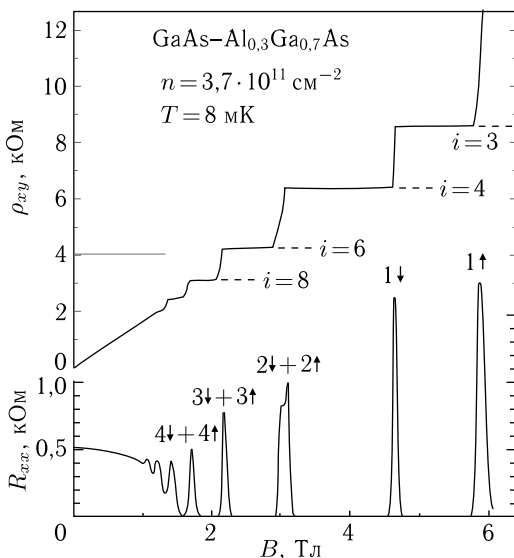


Рис. 9.7. Магнетосопротивление R_{xx} и холловское сопротивление ρ_{xy} в зависимости от магнитного поля B в гетероструктуре $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ [9]. Температура $T = 8$ мК. Электронная плотность в 2D-слое $3,7 \cdot 10^{11} \text{ см}^{-2}$, подвижность $\mu = 4,1 \cdot 10^4 \text{ см}^2/\text{В} \cdot \text{с}$. Стрелками отмечены направления спинов на уровнях Ландау с номером N

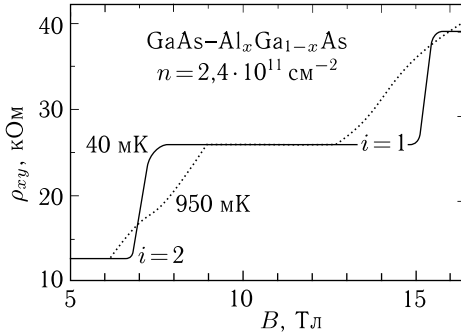


Рис. 9.8. Сужение холловского плато $i = 1$ в гетероструктуре $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ при росте температуры [10]. Электронная плотность в 2D-слое $2,4 \cdot 10^{11} \text{ см}^{-2}$, подвижность $\mu = 9 \cdot 10^4 \text{ см}^2/\text{В} \cdot \text{с}$

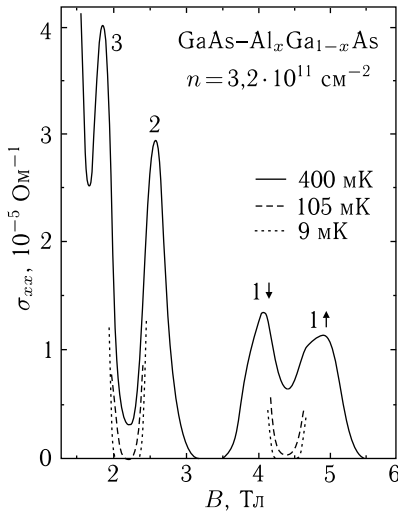


Рис. 9.9. Продольная проводимость в гетероструктуре $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ при разных температурах, измеренная на диске Корбино [11]. Электронная плотность в 2D-слое $3,2 \cdot 10^{11} \text{ см}^{-2}$, подвижность $\mu = 10,5 \cdot 10^4 \text{ см}^2/\text{В} \cdot \text{с}$. Около максимумов указаны самые верхние из полностью заполненных уровней Ландау

временная малость величин $\rho_{xx} \approx 0$ и $\sigma_{xx} \approx 0$ заслуживает особого экспериментального подтверждения. Таким подтверждением является эксперимент в геометрии диска Корбино, непосредственно измеряющий σ_{xx} . Как видно из кривых на рис. 9.9, полученных на диске Корбино, кривая $\sigma_{xx}(B)$ состоит из набора пиков, в промежутках между которыми — в интервалах полей, где находятся холловские плато, проводимость стремится к нулю при понижении температуры. Мы вернемся к температурным зависимостям $R_{xx}(T)$ в параграфе 9.5.

9.3. Механизм образования плато

Начнем обсуждение экспериментальных наблюдений КЭХ, исходя из того, что было изложено в первом параграфе. Представим себе реально неосуществимую *идеальную* 2D-систему без случайного потенциала, в которой мы меняем концентрацию электронов при помощи

напряжения V_g на затворе. На основании уравнений (9.6) и (9.15) можно было бы полагать, что холловская проводимость будет линейно меняться с V_g . Действительно, если бы электронный спектр 2D-газа был непрерывным, то мы бы имели

$$\sigma_{xy} = \frac{ne c}{B} = \frac{c(V_g - V_{g0})}{2\pi s B}, \quad (9.17)$$

где напряжение V_{g0} компенсирует контактные разности потенциалов в цепи. Однако в сильном магнитном поле одна из обкладок конденсатора имеет дискретный энергетический спектр. К чему это приводит, видно из рис. 9.10, где изображена энергетическая диаграмма системы *идеальный 2D-газ плюс затвор* при трех разных напряжениях V_g на затворе.

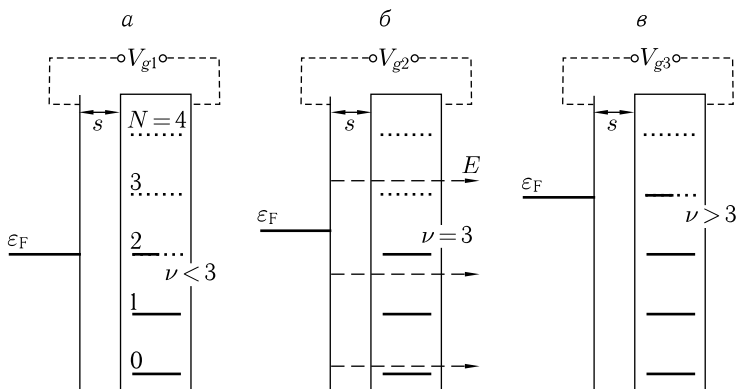


Рис. 9.10. Диаграмма взаимного расположения уровней Ферми в идеальном 2D-газе и в затворе при разных напряжениях V_g на затворе; сплошной линией показаны заполненные состояния, пунктиром — пустые. а) и в) Числа заполнения ν нецелые, а уровень Ферми проходит через уровень Ландау; б) число заполнения ν целое, а плотность состояний на уровне Ферми равна нулю

В равновесии при $V_g - V_{g0} = 0$ уровень Ферми в затворе и в газе находятся на одной высоте, напротив друг друга. При этом положение уровня Ферми относительно уровней Ландау задается исходной концентрацией n_0 . Изменение V_g меняет n . Поэтому наличие разности потенциалов V_g , включенной во внешнюю электрическую цепь, соединяющую затвор с 2D-слоем, будем учитывать смещением уровня ϵ_F в затворе относительно 2D-газа. На диаграмме (а), где напряжение $V_g = V_{g1}$ наименьшее из трех, третий уровень Ландау (считая нулевой уровень) заполнен не полностью и уровень Ферми совпадает с ним. При достижении некоторого значения напряжения третий уровень заполнится, и дальнейшее увеличение концентрации носителей в двумерном газе потребует размещения электронов на четвертом уровне Ландау. Для того чтобы это стало возможным, Ферми-уровень затвора

должен оказаться напротив него, т. е. сдвинуться вверх на $\hbar\Omega$, как это показано на диаграмме (в). А это в свою очередь требует увеличения электрического поля между затвором и 2D-газом. Поэтому в интервале значений напряжения V_g

$$\Delta_{id} = \frac{\hbar\Omega}{e} \quad (9.18)$$

Ферми-уровень будет перемещаться от третьего уровня Ландау к четвертому, находясь при этом в запрещенной энергетической зоне (напряжение V_{g2} на диаграмме (б) относится к этому интервалу напряжений). Двумерный газ при этих V_g является изолятором в том смысле, что он не может сорбировать добавочные электроны и тем самым экранировать внешнее электрическое поле. Дополнительные силовые линии электрического поля, обеспечивающие разность потенциалов $\delta V_g < \Delta_{id}$ проходят через 2D-слой насквозь. При этом концентрация электронов на затворе растет, а в 2D-газе остается постоянной с целым числом заполнения ν . Компенсирующий заряд возникает где-то вне 2D-слоя, в так называемом резервуаре. Роль резервуара могут играть, например, примеси в объеме подложки 2D-слоя, или ее противоположная поверхность, или контакты. Когда при напряжении V_{g3} уровень Ферми достигнет следующего (четвертого) уровня Ландау, заряды из резервуара вернутся в 2D-слой. Концентрация электронов в слое изменится сначала скачком; затем продолжится ее плавное изменение в соответствии с формулой (9.15).

При неизменной концентрации зарядов в 2D-слое холловское напряжение остается постоянным. Поэтому в идеальном материале в зависимости $\rho_{xy}(V_g)$ должны возникнуть плато шириной (9.18) вблизи значений $V_g^{(i)}$, определяемых формулой (9.14). Уровень этих плато $\rho_{xy}^{(i)}$ обусловлен i полностью заполненными уровнями Ландау и поэтому должен иметь точные квантованные значения (9.16).

Следует помнить, что речь идет о мысленном эксперименте в идеальном 2D-слое. Поэтому это рассуждение не может считаться объяснением КЭХ, последний всегда наблюдается на заведомо неидеальных образцах, где часть уровней в 2D-слое локализованы.

Для оценки относительной ширины плато η_{id} сравним величину Δ_{id} с изменением напряжения Δ_ν , требуемым для полного заполнения одного из уровней Ландау, т. е. для изменения ν на единицу, а концентрации на $\Delta n = (2\pi r_B^2)^{-1}$. Магнитное поле, входящее в оба выражения, сократится, и, положив диэлектрическую проницаемость изолирующего слоя конденсатора равной единице, получим

$$\eta_{id} = \frac{\Delta_{id}}{\Delta_\nu} = \frac{1}{s} \frac{\hbar^2}{m e^2} = \frac{a_B}{s}. \quad (9.19)$$

В числителе отношения (9.19) получилось выражение для боровского радиуса a_B . В атоме водорода оно около 0,5 Å. Малая эффективная

масса электронов в 2D-газе может увеличить величину a_B в 10–20 раз, но все равно зазор в конденсаторе s всегда существенно больше. Поэтому для идеального 2D-газа на зависимости $\rho_{xy}(V_g)$ следовало бы ожидать узких и далеко разнесенных друг от друга плато.

Случайный потенциал в 2D-слое уширяет уровни Ландау. Согласно изложенным выше представлениям, набор возникающих при этом состояний включает по крайней мере одно делокализованное состояние с энергией вблизи центра уширенного распределения и локализованные состояния на его периферии. Пусть уширенные уровни Ландау не перекрываются, так что между ними остаются интервалы запрещенных значений энергии. При увеличении V_g Ферми-уровень движется сначала по локализованным состояниям нижней минизоны уровня Ландау, затем по запрещенной зоне, т.е. по области с нулевой плотностью состояний, а затем по локализованным состояниям верхней минизоны (см. рис. 9.11). Заполнение локализованных состояний не должно ска-

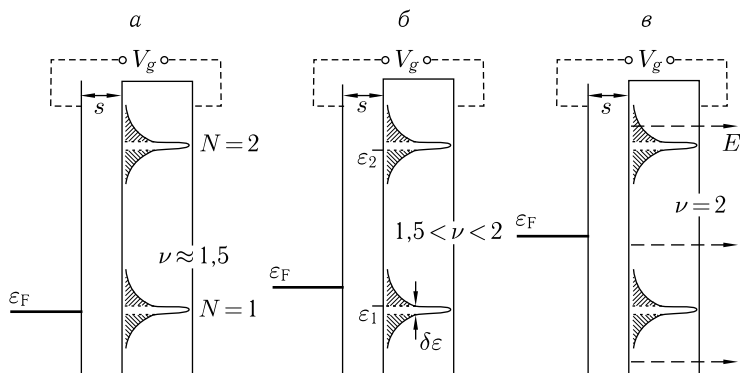


Рис. 9.11. То же, что и на рис. 9.10, но для реального 2D-газа со случайным потенциалом, который превратил уровни Ландау в неперекрывающиеся полосы (минизоны Ландау). Заштрихованы локализованные состояния на краях полос. Нижний уровень $N = 0$ не показан. а) Состояния на Ферми-уровне делокализованы; б) — состояния на Ферми-уровне локализованы; в) плотность состояний на уровне Ферми равна нулю, а число заполнения ν остается целым

зываться на величине эффекта Холла, так что относительная ширина плато η становится

$$\eta = \eta_{\text{id}} + \eta_{\text{loc}} \gg \eta_{\text{id}}, \quad (9.20)$$

где η_{loc} есть доля локализованных состояний на уровне Ландау.

Заметьте. Скорость движения уровня Ферми по энергетическому спектру 2D-газа $d\varepsilon/dV_g$ в областях рис. 9.11, б и рис. 9.11, в различны.

Все сказанное относительно появления плато на кривых $\rho_{xy}(V_g)$ в 2D-системах с затвором справедливо и применительно к кривым

$\rho_{xy}(B)$. Рост V_g эквивалентен уменьшению поля B . Из-за уменьшения вырожденности и количества электронов на нижних уровнях Ландау происходит рост заселенности уровня $\varepsilon(N) = \varepsilon_F$. Рис. 9.12 иллюстрирует эквивалентность изменений V_g и B на примере уширенных уровней Ландау. Поскольку при изменении магнитного поля перетекания зарядов с затвора не происходит, уровень Ферми ε_F следует считать неподвижным; даже если затвора нет, уровень ε_F все равно зафиксирован внешним по отношению к 2D-газу окружением, с которым газ находится в термодинамическом равновесии. При понижении поля «емкость» уровней γ уменьшается (см. формулы (9.8) и (9.9)), и чтобы вместить прежнее количество электронов, система уровней смещается вниз. Пока напротив Ферми-уровня находится крыло уширенного уровня Ландау, происходит заполнение локализованных состояний при неизменном холловском сопротивлении ρ_{xy} (рис. 9.12, б). Когда плотность состояний на Ферми-уровне оказывается равной нулю (рис. 9.12, в), то часть электронов из 2D-слоя уходит в резервуар, и т. д.

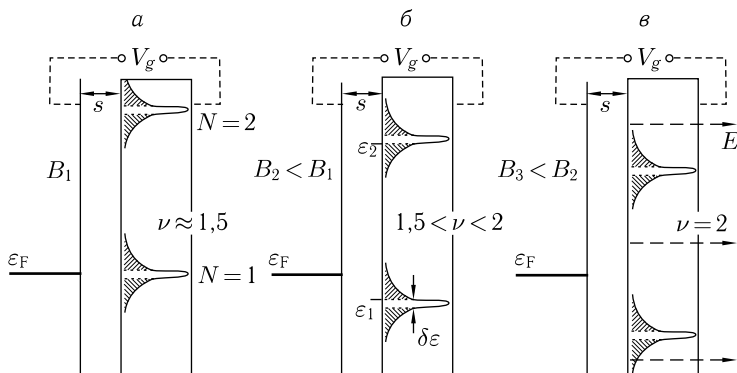


Рис. 9.12. То же, что и на рис. 9.11, но при постоянном V_g и постепенно уменьшающемся магнитном поле B

Объяснив относительное уширение плато за счет локализованных состояний, мы тут же сталкиваемся с основной трудностью в объяснении КЭХ. Локализованные состояния не вносят вклада в холловский ток j_y . В то же время их наличие означает уменьшение числа делокализованных состояний, которые участвуют в создании эффекта Холла. Поэтому основным является вопрос о происхождении *точных* квантовых значений холловского сопротивления на плато. Полагая для определенности, что случайный потенциал $U(\mathbf{r})$ содержит лишь крупномасштабные флуктуации, рассмотрим этот вопрос, следуя идеям работ [12, 13].

На рис. 9.13, а изображена перколяционная сетка эквипотенциальных линий на уровне $U = U_0$. Приложим к 2D-слою электрическое поле $\mathbf{E} \parallel O_x$. Оно снимает вырождение в седловых точках. Энергии точек A_1

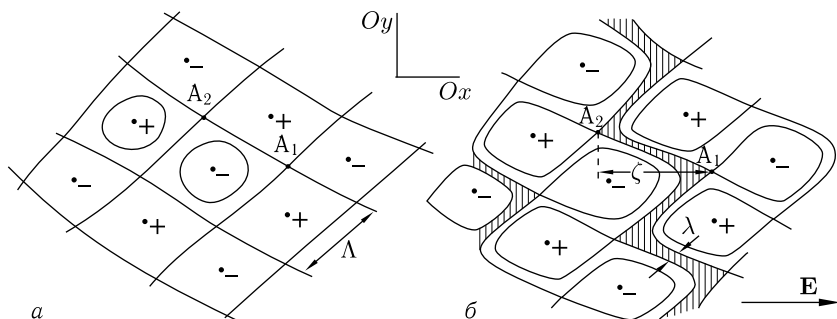


Рис. 9.13. а) Сетка эквипотенциальных линий на уровне перколяции $U = U_0$ в электронном 2D-газе с крупномасштабным случайным потенциалом $U(\mathbf{r})$; б) то же в электрическом поле [12]. Значками $+$ и $-$ отмечены локальные экстремумы функции $U(\mathbf{r})$. Заштрихованы области, из которых эквипотенциали $U = \text{const}$ уходят на бесконечность

и A_2 , ранее принадлежавших одной эквипотенциальной линии, теперь различаются на $eE\zeta$, где ζ — проекция расстояния между точками A_1 и A_2 на направление $\mathbf{E}$. Как видно из рис. 9.13, б, сетка превратилась в систему зигзагообразных полос с шириной λ , зависящей от локального значения градиента случайного потенциала $|\nabla U|_{U_0}$:

$$\lambda(r) = \frac{eE\zeta}{|\nabla U|_{U_0}(r)} \quad (9.21)$$

(здесь r — текущая координата вдоль полосы). Внутри этих полос эквипотенциальные линии уходят на бесконечность вдоль направления Oy , а между полосами имеют форму замкнутых контуров. Поскольку $\lambda(r) \rightarrow 0$ при $E \rightarrow 0$, шириной полосы является большая из величин λ, r_B .

Вернемся к рис. 9.4 и обратим внимание на казалось бы чисто техническую деталь: в выражение для холловского напряжения

$$V_{35} = J_{12}\rho_{xy} \quad (9.22)$$

не входят геометрические размеры a и b ; напряжение V_{35} зависит только от полного тока J_{12} через устройство и от поля B и концентрации n , входящих в ρ_{xy} . Это означает, что правильная геометрическая форма холловского мостика при измерении ρ_{xy} не существенна. Важно лишь, чтобы 2D-газ связывал между собой все задействованные в измерении контакты. Вклады в поперечную проводимость одинаковы от перколяционной сетки уровня Ландау, уширенного случайным потенциалом, и от этого уровня в идеальных условиях полной однородности на плоскости. Если в плоскости, в которой находится идеальный 2D-газ, сделать множество отверстий, то это не отразится на ρ_{xy} , если только не нарушена связность плоскости. Это — центральный пункт для понимания КЭХ.

Каждый уширенный уровень Ландау, расположенный ниже уровня Ферми, содержит внутри себя узкий слой делокализованных состояний; в скрещенных полях этот слой вносит вклад в поперечную проводимость, равный $\Delta\sigma_{xy} = e^2/2\pi\hbar$. Полное значение поперечной проводимости σ_{xy} в режиме плато КЭХ получается умножением $\Delta\sigma_{xy}$ на целое число i таких слоев с делокализованными состояниями ниже уровня Ферми; при этом продольная проводимость отсутствует

$$\sigma_{xx} = 0, \quad \sigma_{xy} = i \left(\frac{e^2}{2\pi\hbar} \right). \quad (9.23)$$

Пусть теперь величины B и V_g таковы, что верхний уровень Ландау заполнен частично (число ν в формуле (9.13) нецелое). При наличии крупномасштабных флуктуаций потенциала занятая 2D-слоем площадь S распадается на две области. В одной верхний уровень Ландау полностью заполнен электронами, в другой — он полностью пустой. Пока дробная часть числа ν мала и в заполненной электронами области 2D-слоя располагаются только замкнутые эквипотенциальные линии конечного размера, т.е. только локализованные состояния, эти электроны не влияют на транспортные свойства. Но как только по мере увеличения концентрации n (или уменьшения поля B) уровень Ферми выйдет на слой делокализованных состояний на перколяционной сетке, контакты холловского мостика на рис. 9.4, *a* окажутся связанными между собой перколяционной сеткой. Эта сетка покрывает лишь часть всей площади, она может состоять из тонких волокон, иметь большие дыры и т.д. — все это не имеет значения. Пока состояния на ней заполнены лишь частично, продольная проводимость σ_{xx} отлична от нуля, а поперечная σ_{xy} имеет неквантованное значение. Когда все состояния в делокализованном слое окажутся полностью заполнены, σ_{xy} выходит на следующее плато.

Уравнение (9.23), связывающее σ_{xy} с числом слоев с делокализованными носителями в энергетическом пространстве под уровнем Ферми, показывает принципиальную важность того, что при уширении каждого уровня Ландау под воздействием возмущающего потенциала $U(\mathbf{r})$ возникает только один слой делокализованных состояний. Уместно задать вопрос: не может ли в пределах одной минизоны Ландау появиться несколько слоев делокализованных состояний с энергиями $\varepsilon_N^{(1)} < \varepsilon_N^{(2)} < \dots$? Такое возможно лишь при очень специфических условиях: если (i) $U(\mathbf{r})$ — периодический потенциал и если (ii) магнитный поток Φ через элементарную ячейку потенциала $U(\mathbf{r})$ отличается от кванта магнитного потока Φ_0 в рациональное число раз, т.е. $\Phi = (p/q)\Phi_0$, где p и q — целые числа. Из второго условия сразу следует, что такой потенциал $U(\mathbf{r})$ должен сам зависеть от магнитного поля.

Заметьте. Речь заведомо не идет о кристаллическом потенциале, на фоне которого всегда реализуются 2D-системы. Кристалличе-

ский потенциал изначально перенормирует электронный спектр, может сделать его многодолинным, анизотропным и т. п., меняет эффективную массу. В иерархии потенциалов он стоит на первом месте. Магнитное поле воздействует уже на перенормированные электроны.

Несмотря на кажущуюся экзотичность, «решеточный вариант» возмущающего потенциала очень важен; отсюда начинается дорога от целочисленного КЭХ к дробному КЭХ. Однако дробный КЭХ не входит в круг обсуждаемых в этой книге вопросов.

9.4. Краевые каналы ¹⁾

В этом параграфе мы обсудим распределение тока в образцах в режиме КЭХ, т. е. при тех значениях B или V_g , при которых реализуются холловские плато с сопротивлением (9.16) и $R_{xx} \ll R_{xy}$. Случайный потенциал будем считать крупномасштабным. Начнем с диска Корбино (рис. 9.4, б), как с более простой измерительной конфигурации: 2D-область в нем не имеет краев, не занятых контактами.

Возникающая из перколяционной сетки система зигзагообразных полос с шириной (9.21) имеет на диске Корбино вид непересекающихся замкнутых линий, охватывающих внутренний электрод. Холловский ток j_y течет вдоль этих линий, огибая области, занятые эквивалентными линиями, не охватывающими внутренний электрод (см. рис. 9.14, заимствованный из [13]). Ток между коаксиальными электродами при сделанных предположениях равен нулю.

Распределение токов в холловском мостике (рис. 9.4, а) сложнее, потому что контакты в нем расположены на краях 2D-области так, что двигаясь вдоль края, можно перейти от одного контакта к другому. Как обсуждалось выше, края 2D-области обладают особыми свойствами. Пусть граница области абсолютно резкая, так что эта область ограничена вертикальными стенками. Из рис. 9.3 видно, что в этом случае вдоль стенок имеется полоска шириной порядка r_D , в которой на Ферми-уровень поднимаются все расположенные ниже уровни Ландау. Она представляет собой одномерную многоканальную транспортную систему с числом каналов равным числу полностью заполненных уровней Ландау, т. е. целой части $i = \text{int}(\nu)$ числа ν (см. описание таких систем в параграфе 5.5).

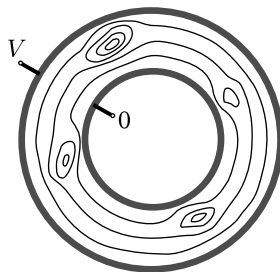


Рис. 9.14. Форма эквивалентных линий в диске Корбино с крупномасштабным случайным потенциалом, находящемся в режиме КЭХ [13]

¹⁾ Смысл и роль краевых каналов подробно обсуждаются в обзоре [14].

Из классической картины движения электронов в этой полоске следует, что электроны сталкиваются с поверхностью на каждом обороте своего движения по ларморовской траектории и, независимо от характера столкновения, смещаются на следующем обороте в том же направлении вдоль канала (рис. 9.1). Главная особенность кинематики

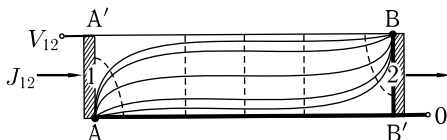


Рис. 9.15. Распределение поля в двухконтактном прямоугольном образце, находящемся в режиме КЭХ при $T = 0$. Потенциал вдоль сплошных линий, соединяющих точки А и В, постояен. Пунктиром показаны линии градиента потенциала, вдоль которых направлено электрическое поле. Плотность эквипотенциальных линий (величина электрического поля) обсуждается в тексте

такого движения в том, что столкновения с поверхностью не приводят к рассеянию назад. Поэтому полоска вдоль края представляет собой не просто 1D-систему, а является *идеальной* многоканальной проволокой типа (5.26), по всем каналам которой при включенном магнитном поле течет конечный ток.

Заметьте. В принципе, если $\nu > 2$, возможно рассеяние из одного канала в другой, хотя при этом должно измениться квантовое число N или направление спина электрона. Однако такое рассеяние не приводит к диссипации. Напротив, уход из канала в объем, который происходит при рассеянии на большой угол на точечной примеси, находящейся на расстоянии порядка r_B от края, приводит к диссипации.

Необходимое уточнение: Предположение о резкой границе 2D-области практически не реализуется. Обычно вертикальной стенки нет, и ширина области, в которой нижние уровни Ландау поднимаются до уровня Ферми, много больше r_B . Поэтому вместо одномерной i -канальной краевой области возникают i одноканальных краевых полосок, идущих параллельно вдоль края 2D-области. При обсуждении реальных экспериментов обычно используется именно такая картина. Лучше всего о ней можно прочитать в оригинальной статье [15]. Для вопроса о распределении холловского тока в 2D-плоскости это уточнение не имеет значения.

Отключим сначала от холловского мостика на рис. 9.4, а контакты 3–6, так что мостик превратится в двухконтактный прямоугольный образец — см. рис. 9.15. Будем считать, что по-прежнему $\rho_{xx} = 0$. Это означает, что хотя вдоль краев А'В и В'А течет ток, диссипации там быть не должно. Следовательно потенциал между контактами вдоль края не меняется. Края А'А и В'В самих контактов, граничащие

с металлом, являются эквипотенциальными линиями по определению. Поэтому прямоугольный контур 2D-области двухконтактного образца распадается на два эквипотенциальных участка $A'B$ и $B'A$, между которыми имеется разность потенциалов V_{12} . В точках A и B происходят скачки потенциала, а все эквипотенциальные линии с потенциалами $0 < V < V_{12}$ тоже собираются к точкам A и B . Конфигурация этих линий при отсутствии случайного потенциала изображена на рис. 9.15. Такая конфигурация потенциального поля кажется удивительной. Однако сделаем небольшое

Лирическое отступление на тему общей физики: Когда мы пускаем ток J по проводу с удельным сопротивлением ρ , укрепленному на стене в виде замысловатого узора, мы задаем Природе трудную задачу: найти решение уравнения Лапласа $\Delta \phi = 0$, удовлетворяющее граничному условию, что электрическое поле $E = -\nabla \phi$ на поверхности провода везде направлено вдоль него и равно $E = J\rho$. Природа решает эту задачу «со скоростью света», располагая дополнительные статические заряды на поверхности провода. Конфигурация поля на рис. 9.15 отнюдь не сложнее. Статические заряды в этом случае располагаются по периметру прямоугольника.

Отношение V_{12}/J_{12} согласно формуле (9.16) равно

$$V_{12}/J_{12} = \frac{1}{i} \left(\frac{2\pi\hbar}{e^2} \right), \quad i = 1, 2, 3, \dots \quad (9.24)$$

Разобьем ток J_{12} на i одинаковых частей J_i по числу зон Ландау под уровнем Ферми и перепишем уравнение (9.24) в виде

$$J_{12} = \sum J_i = i \left(\frac{e^2}{2\pi\hbar} \right) V_{12}. \quad (9.25)$$

Правое равенство (9.25) показывает, что J_i можно считать токами, текущими параллельно по i одномерным идеальным каналам (5.26) вдоль края прямоугольника между контактами 1 и 2. Такой способ описания, сводящий все токи в режиме КЭХ к краевым, связан с именем Бюттикера [14]. Каждая зона Ландау, находящаяся под уровнем Ферми, со своим протяженным состоянием и с подъемом на Ферми уровень по периметру образца, эквивалентна одному такому идеальному каналу.

Обратите внимание на сходство между моделью краевых каналов и статическим скин-эффектом — чисто классическим явлением вытеснения тока на поверхность в чистом металле при низких температурах (см. рис. 9.1).

Возможна и другая формальная трактовка того, как ток J_{12} протекает через образец. Структура выражения (9.25) отражает тот факт, что в холловский ток аддитивно вносят одинаковый вклад i делокализованных слоев всех заполненных уровней Ландау. Вместе с тем из рис. 9.10–9.12 следует, что все эти i слоев находятся ниже уровня

Ферми. Рис. 9.15 и все сказанное в связи с ним относится формально к любому из них. Получается, что холловский ток в 2D-системе может течь по состояниям, лежащим ниже уровня Ферми. Это не противоречит представлениям обычной физики металлов, поскольку холловский ток является недиссипативным. Ток J_{12} с точки зрения внешней задающей цепи, не показанной на рис. 9.15, выполняет роль тока j_x в обозначениях уравнений (9.2), но внутри самого устройства он фактически является током j_y , поскольку везде течет вдоль эквипотенциальных линий электрического поля.

Возьмем один из i заполненных уровней под уровнем Ферми в холловском мостике и изобразим энергию ε_N электрона с соответствующим квантовым числом N как функцию положения центра орбиты. Получится поверхность, по форме напоминающая корыто с дном, «помятым» случайным потенциалом (рис. 9.16). В излагаемой модели ток идет по «дну корыта» вдоль эквипотенциальных линий, изображенных цепочками из точек. Естественно, что окончательный ответ тот же самый, что и в модели краевых токов: между точками А и В течет недиссипативный ток $J_i = (e^2/2\pi\hbar)V_{12}$.

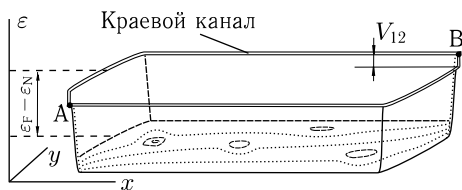


Рис. 9.16. Потенциальная энергия электронов, принадлежащих одному из заполненных уровней Ландау, в зависимости от положения в 2D-области. Точками показаны эквипотенциальные линии, вдоль которых течет недиссипативный ток

Мы изложили две модели распределения холловских токов по 2D-плоскости. Как показал Таулесс [16], на практике реализуется третий, промежуточный вариант. Холловский ток действительно течет вдоль краев 2D-области, но он затухает на гораздо большем расстоянии, чем магнитная длина r_B , причем затухает медленно, как логарифм $\ln r/r_B$ расстояния r до края. Затухание связано с экранированием двумерным электронным газом электрического поля статических зарядов, расположенных по периметру 2D-области и обеспечивающих выполнение граничных условий для уравнения Лапласа. Вследствие экранирования эквипотенциали на рис. 9.15 расположены густо около краев образца и редко в его середине. Поскольку двумерный газ плохо экранирует электрическое поле, плотность силовых линий падает с увеличением r сравнительно медленно.

Таким образом, ток течет частично и вне краевых каналов, и происходит это из-за рассеяния из краевых каналов в объем. Однако он все же сосредоточен вблизи краев, потому что плотность тока пропор-

циональна $|\mathbf{E}| \propto |\nabla U|$ и спадает вглубь из-за экранирования. Поэтому для описания распределения холловских токов можно пользоваться любой из двух моделей. Обычно модель краевых каналов Бюттикера используют применительно к высокоподвижному 2D-газу в широких холловских мостиках, а модель объемных токов — применительно к системам с сильным рассеянием и к узким мостикам.

Вернем теперь холловскому мостику контакты 3–6, т. е. вернемся от конфигурации рис. 9.15 к конфигурации рис. 9.4. Если тока через контакты 3–6 не будет, то по сути дела ничего не изменится, так что $V_1 = V_3 = V_4$ и $V_2 = V_5 = V_6$ и

$$V_{35} = V_{46} = V_{12} = \frac{J_{12}}{i} \left(\frac{2\pi\hbar}{e^2} \right), \quad V_{34} = V_{56} = 0.$$

Из сравнения рис. 9.15 с рис. 9.14 становится понятным, что введение крупномасштабного случайного потенциала ничего не меняет ни качественно, ни даже количественно. Оно сведется к раздвиганию в отдельных местах эквипотенциальных линий АВ и появлению между ними дополнительных замкнутых эквипотенциальных линий. Примечательно, что благодаря краевым каналам несущественно, в каких конкретно местах перколяционная сетка выходит на край 2D-области. Канал вдоль края является эквипотенциальной линией, которая автоматически входит в перколяционную сетку.

Представляет интерес непосредственное доказательство существования токов под уровнем Ферми. Специальные эксперименты на диске Корбино [17] доказали, что перенос заряда под уровнем Ферми действительно происходит. Если не прикладывать извне разность потенциалов к кольцевым электродам 1 и 2 диска Корбино (рис. 9.4, б), а медленно изменять сильное магнитное поле B , нормальное к плоскости диска, то циркулярное и радиальное направления в диске поменяются ролями. В диске возникнет циркулярное электрическое поле $E_{\text{circ}} \propto \partial B / \partial t$, а холловский поток электронов станет радиальным. Поскольку не существует края, связывающего кольцевые электроды, холловский ток может протекать только «через объем», где он вынужден течь под уровнем Ферми. Поставив между контактами 1 и 2 электрометр, этот ток можно зафиксировать.

Заметьте. Механизм возникновения холловского тока $j \propto \sigma_{xy} E_{\text{circ}}$ таков, что электроны не получают энергию от электрического поля E_{circ} . Поэтому во внешней цепи между кольцевыми контактами нельзя использовать для измерения тока обычный амперметр, на внутреннем сопротивлении которого неизбежно диссипируется энергия. С другой стороны, измерить заряд на электродах при помощи электрометра можно только в том случае, если заряд не рассасывается за счет внутренней проводимости σ_{xx} . Это означает, что такая измерительная схема работает только на плато КЭХ, где $\sigma_{xx} = 0$.

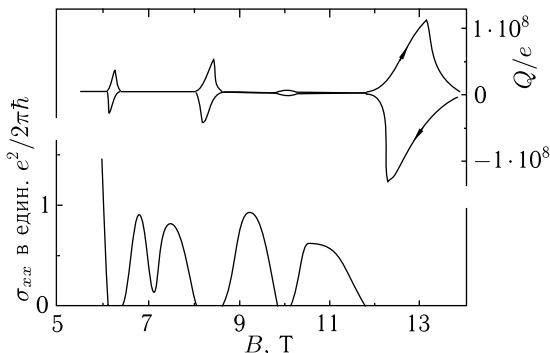


Рис. 9.17. Продольная проводимость σ_{xx} как функция магнитного поля и перенос заряда Q между электродами диска Корбино в процессе изменения поля в режиме КЭХ (направление изменений поля показано стрелками). Диск Корбино с размерами $2a_1 = 225\mu$ и $2a_2 = 675\mu$ изготовлен на базе полевой структуры на поверхности Si. Температура $T = 25$ мК [17]

Результаты такого эксперимента приведены на рис. 9.17. Видно, что заряд Q в электрометре начинает нарастать, как только изменение поля выводит образец на плато КЭХ, где $\sigma_{xx} = 0$. При появлении на противоположном краю плато конечной проводимости σ_{xx} заряд рассасывается. Знак заряда зависит от направления изменения поля. На приведенном графике $Q > 0$ при росте поля и $Q < 0$ при его уменьшении. Скорость нарастания заряда, т. е. наклон прямой $\Delta Q/\Delta B$, позволяет определить величину проводимости. При этом точности обычно недостаточно, чтобы исследовать степень постоянства величины σ_{xy} вдоль плато КЭХ, но вполне достаточно, чтобы различить целые числа i в уравнении (9.23).

9.5. Плотность состояний электронного 2D-газа в магнитном поле

Плотность состояний, схематически изображенная на рис. 9.11 и рис. 9.12, представляет фундаментальный интерес и является целью многочисленных экспериментальных исследований. Остановимся сначала на измерениях плотности локализованных состояний.

Плато КЭХ реализуется при тех значениях V_g или B , когда уровень Ферми находится в области запрещенных или локализованных состояний на расстояниях ε_1 и ε_2 от ближайших уровней делокализованных состояний. При этом диссипативная проводимость реализуется за счет термического заброса носителей на эти уровни, выполняющих роль края подвижности. Это означает, что следует ожидать активационного закона изменения продольного сопротивления и продольной проводимости

$$\rho_{xx} \propto \exp\left(-\frac{\varepsilon^*}{T}\right), \quad \sigma_{xx} \propto \exp\left(-\frac{\varepsilon^*}{T}\right). \quad (9.26)$$

На рис. 9.18 приведены результаты эксперимента на холловском мостике на гетероструктуре $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$. Сопротивление R_{xx} действительно меняется по активационному закону (9.26), причем величина ε^* , определяемая из наклона прямых на рис. 9.18, *а*, сильно меняется с магнитным полем. Происходящие при этом физические процессы можно пояснить при помощи рис. 9.11 и 9.12. Продольное сопротивление обратно пропорционально концентрации носителей на уровне $N = 2$ или дырок на уровне $N = 1$. Поэтому величина ε^* , входящая в соотношение (9.26), равна

$$\varepsilon^* = \min(|\varepsilon_F - \varepsilon_1|, |\varepsilon_F - \varepsilon_2|), \quad \varepsilon_2 - \varepsilon_1 \approx \hbar\Omega. \quad (9.27)$$

Последнее соотношение выполняется с точностью до отношения $\delta\varepsilon/\hbar\Omega$ ширины слоя делокализованных состояний к циклотронной щели.

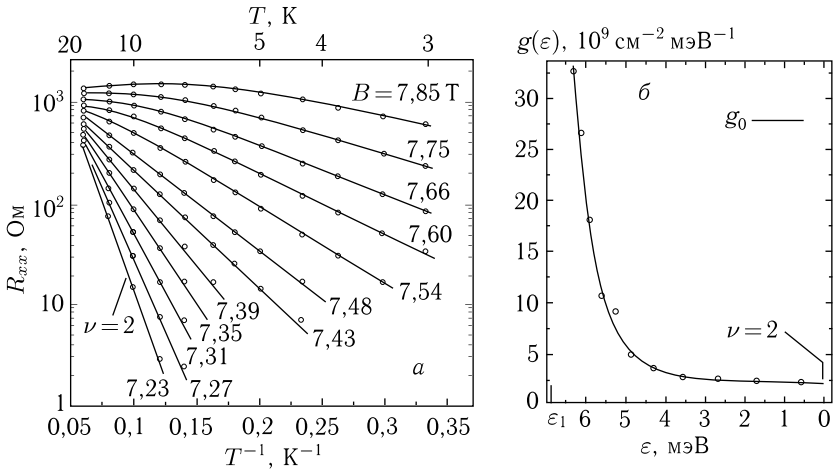


Рис. 9.18. *а*) Температурная зависимость продольной проводимости гетероструктуры $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ в интервале температур 20–3 К при различных значениях магнитного поля. Подвижность $\mu = 5,5 \cdot 10^5 \text{ cm}^2/\text{B} \cdot \text{с}$ [18]. *б*) Плотность локализованных состояний в щели между уровнями Ландау, построенная на основании этого эксперимента [18]. ε_1 — середина уширенного уровня Ландау $N = 1$, g_0 — плотность состояний без магнитного поля

Наклон прямой $\ln R_{xx}$ от $1/T$ максимален, когда уровень Ферми находится точно посередине между двумя краями подвижности. Это соответствует целому значению числа заполнения ν . На рис. 9.18 максимальный наклон достигается при $\nu = 2$. Изменение наклона $\Delta\varepsilon^*$ при изменении B или V_g определяет сдвиг уровня Ферми $\Delta\varepsilon_F = \Delta\varepsilon^*$. Число электронов Δn , которое при этом размещается на локализованных уровнях в интервале энергий $\Delta\varepsilon_F$, определяется формулой (9.15) из ΔV_g либо формулой (9.13) из дробной части числа ν . Отсюда сразу

следует плотность состояний в энергетическом зазоре между уровнями Ландау $g(\varepsilon) = \Delta n / \Delta \varepsilon$. Как видно из рис. 9.18, б, она оказалась неожиданно высокой, около 10 % от плотности состояний g_0 в нулевом магнитном поле.

Однако заметьте: В измеряемую плотность состояний входят электроны, размещаемые в резервуаре. Это особенно существенно, если локализованные уровни резервуара размещаются близко от 2D-газа, так что между ними устанавливается термодинамическое равновесие.

Наличие конечной плотности состояний в щели между уровнями Ландау означает, что в режиме КЭХ при достаточно глубоком понижении температуры проводимость (9.26) по делокализованным состояниям должна смениться прыжковой проводимостью с переменной длиной прыжка (см. гл. 4) по состояниям в непосредственной окрестности уровня Ферми. Рис. 9.19 и 9.20, где температурный интервал ниже, чем на рис. 9.18, показывают, что такая смена режима продольной проводимости действительно происходит. Степень $T^{-1/3}$ в показателе активационной экспоненты указывает на постоянную плотность состояний в окрестности уровня Ферми, а степень $T^{-1/2}$ обычно является

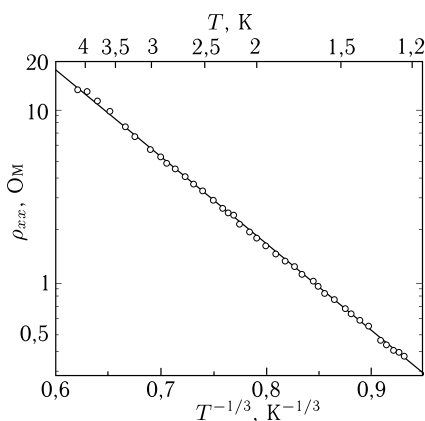


Рис. 9.19. Температурная зависимость продольного сопротивления гетероструктуры GaAs-Al_xGa_{1-x}As в интервале температур 5–1 К в режиме КЭХ. Электронная концентрация $n = 2,4 \cdot 10^{11} \text{ см}^{-2}$, подвижность $\mu = 0,8 \cdot 10^5 \text{ см}^2/\text{В} \cdot \text{с}$ [19]

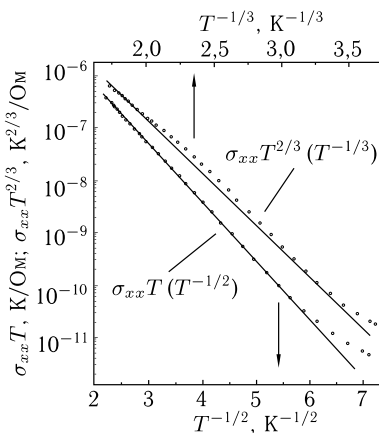


Рис. 9.20. Температурная зависимость продольной проводимости гетероструктуры GaAs-Al_xGa_{1-x}As в геометрии Корбино в интервале температур 0,2–0,02 К в режиме КЭХ при $B = 4,4 \text{ Тл}$. Электронная концентрация $n = 3,2 \cdot 10^{11} \text{ см}^{-2}$, подвижность $\mu = 0,38 \cdot 10^5 \text{ см}^2/\text{В} \cdot \text{с}$ [11]. Оригинальные кривые $\sigma_{xx}(B)$ на этом образце при разных температурах см. рис. 9.9

индикатором наличия кулоновской щели. То, что в двух экспериментах получились две разные степени, может быть результатом смены режима прыжковой проводимости при смещении вниз температурного интервала — этот интервал на рис. 9.19 расположен в области более низких температур, чем на рис. 9.20 (ср. рис. 3 в гл. 4). Возможно однако, что это различие просто отражает разницу в свойствах образцов.

Более интересным и в каком-то смысле более принципиальным является вопрос о плотности состояний в окрестности делокализованных уровней. Вернемся к рис. 9.13 (а), на котором изображена перколяционная сетка при крупномасштабном случайном потенциале $U(\mathbf{r})$. Формально случайный потенциал полностью снимает вырождение, и уровень, соответствующий перколяционной сетке, невырожден. Но мы уже упоминали о различных факторах, которые способствуют тому, что электроны на очень близких по энергии уровнях тоже фактически делокализованы. К ним относятся магнитный пробой, случайные короткопериодные флуктуации, конечные размеры образцов, конечная температура и т. п. Чтобы оценить количество делокализованных состояний в одной зоне Ландау, предположим, что характерный размер ячеек перколяционной сетки равен Λ . Этот размер определяется природой потенциала. Например, если случайный потенциал $U(\mathbf{r})$ создается примесями, расположенными в глубине слое изолятора в некотором отдалении от 2D-слоя, то $\Lambda \propto \mathcal{N}^{-1/2}$, где $\mathcal{N}$ — это 2D-концентрация этих примесей. Общая длина «волокон» этой сетки на единицу площади 2D-слоя порядка $2/\Lambda$. Считая, что ширина «волокна» не меньше, чем $2r_B$, получим, что волокна бесконечной сетки покрывают собой по крайней мере долю $4r_B/\Lambda$ от всей площади 2D-слоя. Из формулы (9.8) следует, что и доля делокализованных состояний η_{del} порядка

$$\eta_{\text{del}} \sim r_B/\Lambda, \quad (9.28)$$

величина η_{del} есть доля делокализованных состояний в минизоне Ландау.

9.6. Цепочки фазовых переходов

При взгляде на кривые $R_{xx}(B)$ и ρ_{xy} на рис. 9.7 и 9.8, а также σ_{xx} на рис. 9.9 сразу возникает мысль о том, что состояния на плато КЭХ являются особыми фазовыми состояниями со специфическими транспортными свойствами

$$\sigma_{xx} \rightarrow 0, \quad \sigma_{xy} = i(e^2/2\pi\hbar), \quad i = 0, 1, 2, 3, \dots \quad (9.29)$$

Фазы (9.29) с $i > 0$ нельзя считать фазами изолятора из-за свойств $\rho_{xx} = 0$ и $\sigma_{xy} = i(e^2/2\pi\hbar) \neq 0$. В то же время свойство $\sigma_{xx} = 0$ не позволяет считать их металлическими. Поэтому их называют [20] квантовыми холловскими жидкостями с разными квантовыми холловскими числами. Какое квантовое холловское число i реализуется на

практике, зависит от числа заполнения (9.13). Чтобы разобраться, как происходит смена квантового числа i на $i \pm 1$ при плавном изменении ν , обратимся к рис. 9.11 или рис. 9.12, где параметром, управляющим числом заполнения ν , являются напряжение V_g или поле B . Пусть число заполнения ν таково, что уровень Ферми попадает на уровень делокализованных, протяженных, состояний. Если интегральная плотность n_{del} протяженных состояний сравнительно невелика, то уровень Ферми при изменении ν проходит соответствующую область быстро, и при изменении управляющего параметра происходит фазовый переход сразу из одного состояния квантовой холловской жидкости в другое. Этот переход сопровождается скачком поперечной проводимости на $|\Delta\sigma_{xy}| = e^2/2\pi\hbar$. Металлическими свойствами при этом обладает лишь пограничное состояние, но с обеих сторон от перехода $\sigma_{xx} = \rho_{xx} = 0$. Если же плотность состояний в окрестности этого уровня велика, то слой делокализованных состояний закрепляет уровень Ферми. Тогда фазовый переход расщепляется на два, и между двумя состояниями холловской жидкости со значениями i , отличающимися на единицу, появляется металлическое состояние, при котором на уровне Ферми находится частично заполненный слой протяженных состояний. Какой из двух вариантов осуществляется, зависит от множества факторов, определяющих конкретную реализацию случайного потенциала.

Все сказанное можно проиллюстрировать конкретными экспериментальными кривыми. При наличии промежуточной металлической фазы наклон между ступеньками ρ_{xy} и ширина пика R_{xx} стремятся при понижении температуры к конечным значениям. Именно так вели себя кривые на рис. 9.7, причём пики R_{xx} росли с понижением температуры [9]. Кривые на рис. 9.21, взятом из работы [21], выглядят иначе: пики ρ_{xx} между плато ρ_{xy} с малыми i уменьшаются и пропадают при низких температурах. При 50 мК пик $1 \uparrow$ уже вовсе отсутствует. И это несмотря на то, что параметры образцов в обоих экспериментах очень близки.

Когда наклон $\partial\rho_{xy}/\partial\nu$ перемычки между двумя ступенями стремится при понижении температуры к конечному значению, то из него можно оценить, какая доля (9.28) состояний на уровне Ландау принадлежит слою делокализованных электронов. В образце, кривые с которого представлены на рис. 9.7, эта доля составляет для нижних уровней около 3%.

Область, соответствующая квантовому числу $i = 0$, должна обладать особыми свойствами: согласно формуле (9.23), в ней равны нулю все компоненты тензора $\hat{\sigma}$: $\sigma_{xx} = \sigma_{xy} = 0$, а диагональные компоненты тензора $\hat{\rho}$ обращаются в бесконечность: $\rho_{xx} = \infty$. Поэтому для состояний в этой области было даже предложено особое название: «холловский изолятор».

Энергетическая диаграмма на рис. 9.11 или рис. 9.12 с симметрично-уширенными уровнями Ландау предполагает, что переходы проис-

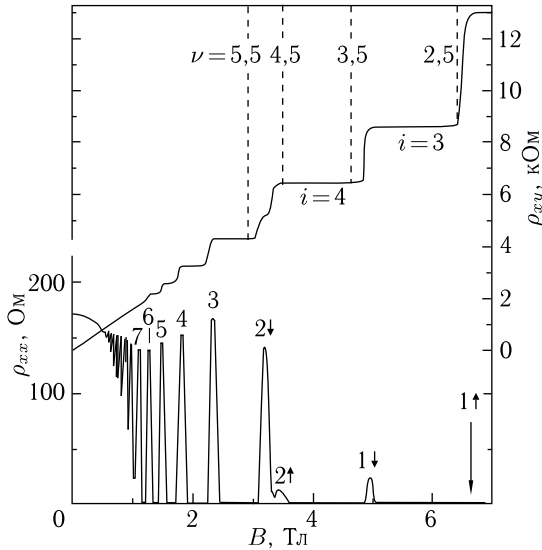


Рис. 9.21. Магнетосопротивление ρ_{xx} и холловское сопротивление ρ_{xy} гетероструктуры GaAs-Al_xGa_{1-x}As при температуре 50 мК. Электронная концентрация $n = 4 \cdot 10^{11} \text{ см}^{-2}$, подвижность $\mu = 0,86 \cdot 10^5 \text{ см}^2/\text{В} \cdot \text{с}$ [21]. На кривой ρ_{xx} пик $1\downarrow$ сравнительно маленький, а пик $1\uparrow$ и вовсе отсутствует. Вертикальными пунктирными линиями показаны полуцелые числа заполнения ν , при которых в простейшей модели должны были бы происходить фазовые переходы со скачками холловского сопротивления ρ_{xy} (вертикальные линии взяты из книги [1])

ходят при полуцелых значениях чисел заполнения

$$\nu = \nu_c^0 = i + 1/2, \quad i = 0, 1, 2, \dots \quad (9.30)$$

На рис. 9.21, где представлены результаты экспериментов при фиксированной концентрации n , эти уравнения определяют значения B_i полей, в которых ожидаются фазовые переходы и соответствующие ступеньки на кривой ρ_{xy} . Как видно из рис. 9.21, соотношение (9.30) выполняется лишь весьма приблизительно.

Можно указать несколько возможных причин отклонений значений ν_c от ν_c^0 .

Во-первых, это асимметрия в плотности и расположении положительно и отрицательно заряженных источников случайного потенциала, которая смещает вырожденный уровень протяженных состояний из центра зоны Ландау. Во-вторых, причиной сдвига значений ν_c может быть превышение уширения уровней над расстоянием между ними. Если хвост локализованных состояний одного уровня ($\mathcal{A}$) выходит за максимум плотности состояний расположенного ниже уровня $\mathcal{B}$, то прежде чем уровень Ферми достигнет максимума на уровне $\mathcal{B}$, должны будут заполниться часть состояний в хвосте уровня $\mathcal{A}$. Это приводит

к сближению двух переходов на шкале ν из начальных положений $i - 1/2$ и $i + 1/2$. Именно такое сближение переходов $2\uparrow$ и $2\downarrow$, по-видимому, наблюдается на рис. 9.21.

В малых магнитных полях существует еще одна возможная причина сдвига фазовых границ в сторону больших ν , получившая название *всплывание уровней*. Всплывание было предсказано Хмельницким [24] на основании того, что описание КЭХ с делокализованными состояниями в середине каждого уровня Ландау должно быть согласовано, при предельном переходе $B \rightarrow 0$, со скейлинговой гипотезой, утверждающей, что в нулевом магнитном поле все электронные 2D-состояния должны быть локализованными (см. гл. 6). Если уменьшать магнитное поле, то все большее число уровней Ландау уходят под уровень Ферми: $N \approx \varepsilon_F / \hbar \Omega$. Вместе с ними должно уходить вниз и такое же число протяженных состояний. Протяженные состояния не могут постепенно превратиться в локализованные, между тем в нулевом поле их не должно быть.

Для разрешения кажущегося парадокса требуется проанализировать транспортные свойства 2D-электронного газа в области $\Omega\tau < 1$.

К вопросу о терминологии: Поля $\Omega\tau < 1$ обычно называют слабыми. Однако в данной главе речь идет о полях $\hbar\Omega(N_0 + 1/2) \gtrsim \varepsilon_F$, где $N_0 = 4 \div 5$ (см. рис. 9.2), т.е. о полях, в которых магнитная длина (9.9) лишь чуть больше среднего расстояния $n^{-1/2}$ между 2D-электронами, $B \sim (c\hbar/e)n$. Поэтому здесь неравенство $\Omega\tau < 1$ означает «сильное поле при очень сильном, $\hbar/\tau \sim \varepsilon_F$, беспорядке».

Возьмем классические выражения (9.5) для компонент тензора проводимости в магнитном поле и следующее из них соотношение

$$\sigma_{xy} = (\Omega\tau)\sigma_{xx}. \quad (9.31)$$

Они должны работать локально, на малых масштабах порядка магнитной длины r_B , определенной формулой (9.9). На масштабах порядка размера образца справедливы формулы КЭХ, которые дают для поперечной проводимости σ_{xy} на плато значения (9.23), а в промежутках между плато, когда на уровень Ферми выходит $(i + 1)$ -е делокализованное состояние, значение

$$\sigma_{xy} = (i + 1/2) \left(\frac{e^2}{2\pi\hbar} \right). \quad (9.32)$$

Для того чтобы сшить эти два масштаба, подставим в (9.31) квантовое выражение (9.32) для σ_{xy} , оставив для σ_{xx} классическое выражение (9.5). Разрешив получившееся уравнение относительно концентрации электронов n , входящей в классическое выражение для σ_{xx} , получим

$$n = (i + 1/2) \frac{1}{2\pi r_B^2} \frac{1 + (\Omega\tau)^2}{(\Omega\tau)^2}. \quad (9.33)$$

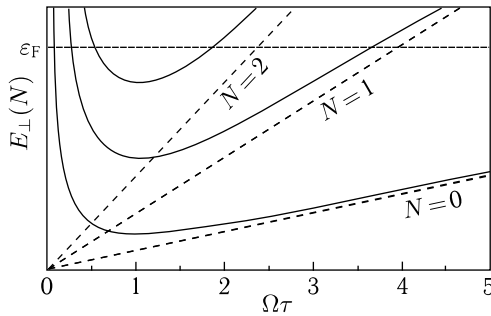


Рис. 9.22. Всплывание уровней. Сплошные линии — энергия протяженных состояний в зависимости от $\Omega\tau$ в соответствии с формулой (9.34); пунктирные прямые, выходящие из начала координат, показывают положение уровней Ландау (9.30), заполненных наполовину

Поскольку в 2D-газе концентрация n линейно связана с энергией Ферми, из (9.33) получается выражение для энергии N -того протяженного состояния (см. рис. 9.22)

$$E_{\perp}(N) = (N + 1/2)\hbar\Omega \frac{1 + (\Omega\tau)^2}{(\Omega\tau)^2} = \varepsilon_{\perp}(N) (1 + 1/(\Omega\tau)^2), \quad (9.34)$$

в котором для удобства сравнения использованы те же обозначения, что и в выражении (9.7) в начале главы ($E_{\perp}$, $\varepsilon_{\perp}$, N вместо i).

При выводе уравнений (9.33) и (9.34) был использован отнюдь не бесспорный прием, позволивший сравнить проводимости на разных масштабах. Дополним его довольно произвольным обобщающим предположением, что формула (9.34) сохраняет смысл при $\Omega\tau < 1$. Поскольку дробь, входящая в уравнение (9.34), порядка единицы при больших $\Omega\tau$ и быстро растет при малых $\Omega\tau$, энергия $E_{\perp}(N)$ растет при уменьшении $\Omega\tau$. Это и называется всплыванием протяженных уровней. При понижении магнитного поля энергия протяженных уровней растет, и они один за другим пересекают уровень Ферми.

Заметьте. Рост функции (9.34) в малых полях, т.е. всплывание, о котором идет речь, происходит в области $\Omega\tau < 1$, где уже плохо работает сама классификация уровней по Ландау. Поэтому выражение (9.34) не может рассматриваться как зависимость от B энергии конкретного протяженного уровня. Но существование протяженных уровней и их энергия не зависят от того, в каком базисе мы пытаемся их описать. Поэтому можно надеяться, что скейлинговая процедура, использованная при переходе от (9.31) к (9.33), правильно определяет тенденцию и знак квантовой поправки, обусловленной рассеянием. Задача состоит в экспериментальном подтверждении этой тенденции.

Заметьте также, что формула (9.34) относится именно к энергии протяженных состояний, а не к центру тяжести уширенного уровня Ландау.

Рассуждения, на основании которых был сделан вывод о всплывании уровней, нельзя считать достаточно строгими. По сути дела, это есть распространение скейлинговой гипотезы (гл. 6) на двумерные системы в магнитном поле. Попытки сделать рассуждения более строгими сразу же приводят к серьезному усложнению математического аппарата (см., например, [22]). Поэтому, как и в гл. 6, мы сосредоточимся на экспериментальной проверке сделанных выводов.

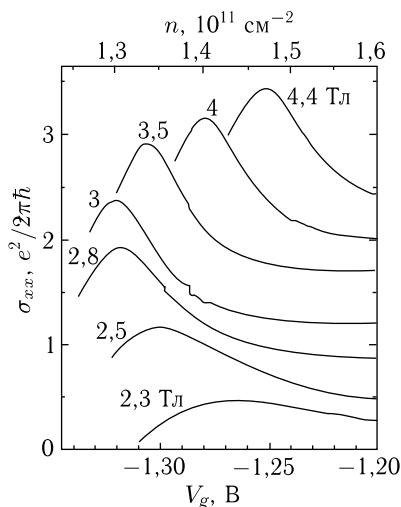


Рис. 9.23. Максимум σ_{xx} , соответствующий нижнему делокализованному состоянию в гетероструктуре GaAs–Al_xGa_{1-x}As с низкой подвижностью, при разных значениях электронной плотности [23]. Шкала слева соответствует кривой в поле 3 Тл; остальные кривые сдвинуты для ясности вверх (кривые в больших полях) или вниз (кривые в меньших полях). Температура $T = 25$ мК

Эксперимент подтвердил идею о всплывании уровней, по крайней мере частично. Лучше всего всплывание видно при измерениях положения максимумов диссипативной компоненты сопротивления ρ_{xx} при изменении поля B или концентрации $n \propto V_g$. Максимумы наблюдаются при тех значениях B или n , когда протяженное состояние оказывается на уровне Ферми. В простейшей модели это прямые (9.30), исходящие из начала координат на плоскости (B, n) , так что концентрация должна понижаться линейно с полем. То, что при малых i это может быть не так, видно из рис. 9.23, взятом из работы [23]: при понижении поля максимум функции $\sigma_{xx}(n)$ сначала действительно движется в сторону меньших концентраций, но затем начинает смещаться обратно. Из рисунка ясны также пределы возможностей этого эксперимента по понижению поля: в полях $\Omega t < 1$ максимум расширяется, а точность измерений падает. Тем не менее, увеличение энергии протяженных состояний в малых полях удастся наблюдать абсолютно надежно (рис. 9.24).

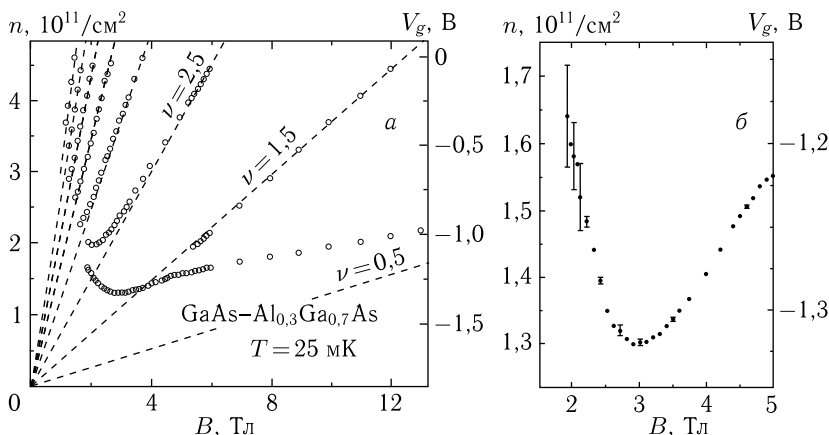


Рис. 9.24. а) Положение нескольких нижних делокализованных состояний в гетероструктуре $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ с низкой подвижностью на плоскости (B, n) [23]. Пунктирные прямые проведены в соответствии с уравнением (9.30). б) То же для самого нижнего состояния в увеличенном масштабе [23]

Уравнение (9.34) и, соответственно, рис. 9.22 предполагают, что протяженные делокализованные состояния сохраняют свою индивидуальность вплоть до пересечения ферми-уровня. Это предсказание теории до сих пор не удалось ни подтвердить, ни опровергнуть. Из многих работ следует, что энергии протяженных состояний не опускаются ниже некоторого уровня, и что энергия самого низкого из них растет при понижении поля. Однако остается неизвестным, как они ведут себя по отношению друг к другу.

9.7. Двухпараметрический скейлинг ¹⁾

В гл. 6 обсуждалась скейлинговая гипотеза в приложении к переходу металл–изолятор, контролируемому беспорядком. Задача состоит в том, чтобы обобщить использованный там подход, применив его к 2D-системам в сильном магнитном поле. Мы уже начали этот процесс, сформулировав и проверив предположение о всплывании уровней. Теперь предстоит сделать следующий шаг, нарисовав что-то вместо функций $\beta(y)$ на рис. 6.1.

В отличие от описания перехода металл–изолятор, в 2D-системе при выборе функции, задающей состояние системы, не нужно выбирать между удельной проводимостью и кондактансом, поскольку они совпадают. Зато из-за магнитного поля проводимость имеет теперь две

¹⁾ Для глубокого понимания идей двухпараметрического скейлинга полезно прочесть дополнительно обзор [25] и короткую статью того же автора [26]

компоненты: σ_{xx} и σ_{xy} , и для каждой можно написать логарифмическую производную проводимости σ по размеру системы L в виде функции проводимости (как и при написании уравнения (6.3) в гл. 6 предполагается, что температура $T = 0$):

$$\begin{aligned}\frac{d \ln \sigma_{xx}}{d \ln L} &= f_1(\sigma_{xx}, \sigma_{xy}), \\ \frac{d \ln \sigma_{xy}}{d \ln L} &= f_2(\sigma_{xx}, \sigma_{xy}).\end{aligned}\tag{9.35}$$

Если при переходе металл–изолятор поведение системы определялось одним параметром: кондактансом y , то здесь таких независимых параметров два: σ_{xx} и σ_{xy} . Отсюда название этого параграфа.

Исключив переменную L из двух уравнений (9.35), получим связь между σ_{xx} и σ_{xy} , которую можно изобразить в виде кривых на плоскости $(\sigma_{xy}, \sigma_{xx})$ [27]. Эта плоскость называется фазовой, а сами кривые — фазовыми траекториями.

К вопросу о терминологии:

Подобное представление широко используется в механике, где движение системы описывают при помощи траектории, изображающей точки на плоскости $(q, \dot{q})$, где q — это обобщенная координата. Отсюда термины «фазовая плоскость» и «фазовая траектория». В них под словом фаза подразумевается характеристика движения, а не состояние вещества. Поскольку в обсуждаемой нами проблеме центральным является именно вопрос о состоянии вещества, то во избежание недоразумений совокупность кривых, описывающих двухпараметрический скейлинг, часто называют диаграммой потока. Этот термин также заимствован из механики. Однако, как мы увидим ниже, диаграмма потока содержит также линии, определяющие положение межфазовых границ на плоскости $(\sigma_{xy}, \sigma_{xx})$, и поэтому является одновременно фазовой диаграммой, определяющей возможные переходы между различными состояниями 2D-электронов.

Фазовая траектория механической системы, проходящая через заданную начальными условиями точку $(q_0, \dot{q}_0)$, определяет эволюцию этой системы со временем, которое является скрытым параметром. На фазовой плоскости КЭХ роль времени играет размер L , а под начальными условиями следует понимать значения компонент тензора проводимости (9.5) в классическом пределе, справедливые на размерах L порядка магнитной длины (9.9). Движение по фазовой траектории означает постепенный переход к большим размерам L . Сначала это должно привести к включению квантовых поправок. Затем рост L приведет к тому, что поправки станут большими и начнутся процессы локализации, в результате чего система перейдет в состояние КЭХ.

Пользуясь этими соображениями, представим на рис. 9.25, а «эскизный проект» плоскости со скейлинговыми фазовыми траекториями. Это должен быть вариант скейлинговой гипотезы, пригодный для

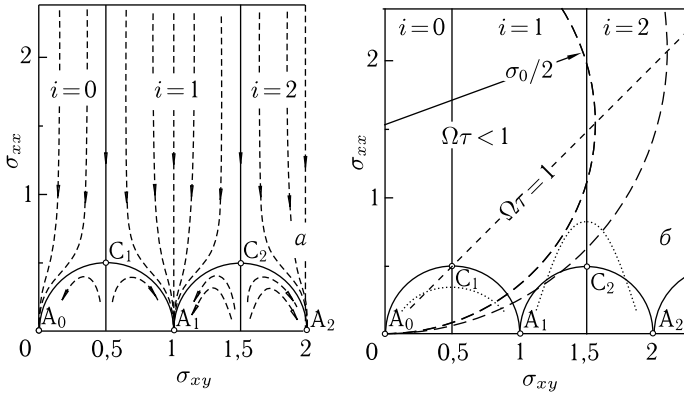


Рис. 9.25. а) Скейлинговая фазовая плоскость, иначе называемая диаграммой потока; координатами являются компоненты тензора проводимости σ_{xy} и σ_{xx} в безразмерных единицах $e^2/2\pi\hbar$. Сплошные линии — сепаратрисы, A_i — стационарные особые точки, C_i — нестационарные особые точки. б) Положение начальных точек фазовых траекторий, определяемых значениями n, τ и B , относительно сепаратрис диаграммы потока. Пунктирная прямая под наклоном 45° разделяет области $\Omega\tau > 1$ (внизу) и $\Omega\tau < 1$ (вверху). Две пунктирные полуокружности — геометрические места (9.37) точек, удовлетворяющих уравнениям (9.5) при фиксированных σ_0 и переменном $\Omega\tau$. Относительно кривых, показанных точками, см. в тексте

2D-систем в сильном магнитном поле. При $L \rightarrow \infty$ реализуются КЭХ и проводимости принимают значения (9.23), так что *все* траектории должны собраться в точки

$$((e^2/2\pi\hbar)i, 0), \quad i = 0, 1, 2, \dots \quad (9.36)$$

на оси σ_{xy} . На рис. 9.25 это точки A_i . Они являются стационарными особыми точками фазовой диаграммы: все траектории в окрестности такой точки направлены к ней. Каждая точка A_i соответствует своей фазе квантовой холловской жидкости с квантовым числом $i > 0$ (здесь слово фаза уже имеет смысл состояния вещества), либо холловскому изолятору при $i = 0$. Вся фазовая плоскость разбивается на области, внутри каждой из которых все траектории собираются в одну из точек A . Линии, разделяющие эти области, называются сепаратрисами. Выберем в качестве начальной некоторую точку $(\sigma_{xy}^0, \sigma_{xx}^0)$. Квантовые поправки к продольной проводимости подробно обсуждались в гл. 2. Поскольку здесь речь идет о сильных магнитных полях, слабую локализацию можно считать уже разрушенной, так что остается межэлектронное взаимодействие. Оно приводит к логарифмической поправке (2.37) к σ_{xx} и к нулевой поправке к σ_{xy} : в первом приближении теории возмущений $\Delta\sigma_{xy} = 0$. Поэтому начальные участки траекторий потока — это вертикальные прямые линии. Соответственно вертикальными прямыми линиями являются и сепаратрисы.

В каждом промежутке значений $i < \sigma_{xy}/(e^2/2\pi\hbar) < i + 1$ между стационарными точками (9.36) при некотором $\sigma_{xx} = \sigma_{xx}^* > 0$ имеется нестационарная особая точка С, вблизи которой поток траекторий разветвляется. К сказанному следует прибавить требование периодичности по оси σ_{xy} с периодом $e^2/2\pi\hbar$, поскольку квантовые фазы с разными числами i эквивалентны. Тогда система сепаратрис приобретает вид, показанный на рис. 9.25 сплошными линиями. При этом вертикальные сепаратрисы одновременно разделяют области, в которых реализуются разные квантовые состояния, и поэтому являются межфазовыми границами, а вся диаграмма потока является одновременно и диаграммой фазовых состояний. Самая левая из межфазовых границ, проходящая через точку С₁, отделяет состояние изолятора $i = 0$ от состояния квантовой холловской жидкости $i = 1$. Согласно диаграмме, с другими состояниями холловской жидкости, $i > 1$, изолятор не граничит. Это означает, что возможны только переходы $0 \rightarrow 1$, а переходы $0 \rightarrow 2$, $0 \rightarrow 3$ и т. д. невозможны. Это одно из основных утверждений, подлежащих экспериментальной проверке.

Начальной точкой $(\sigma_{xy}^0, \sigma_{xx}^0)$, обусловленной классической проводимостью 2D-системы, реализующейся на малых масштабах, может быть практически любая точка плоскости $(\sigma_{xy}, \sigma_{xx})$, кроме точек в узкой полосе вдоль оси абсцисс, где расположены фазовые траектории, описывающие дробный КЭХ. (Эту область мы обсуждать не будем; интересующимся можно порекомендовать фундаментальную работу [20]). Расположение начальных точек поясняет рис. 9.25, б. Координаты этих точек зависят от трех параметров: концентрация n , времени упругой релаксации τ и магнитного поля B . Произведение двух первых определяет проводимость материала в нулевом поле σ_0 , два последних входят в безразмерное произведение $\Omega\tau$. Если зафиксировать n и τ , положив тем самым $\sigma_0 = ne^2\tau/m = \text{const}$, то получим из соотношений (9.5), что при изменении магнитного поля $0 \leq B \leq \infty$ изображающая точка на плоскости очерчивает полуокружность

$$(\sigma_{xy}^0)^2 + (\sigma_{xx}^0 - \sigma_0/2)^2 = \sigma_0^2/4. \quad (9.37)$$

Ее верхняя часть соответствует полям $\Omega\tau < 1$, а нижняя часть — полям $\Omega\tau > 1$.

Верхнюю часть полуокружности (9.37) можно в принципе считать геометрическим местом начальных точек фазовых траекторий. В пределе $\Omega\tau \gg 1$ при вычислении классической проводимости σ_{xx}^0 из кинетического уравнения существенно квантование электронного спектра, не учтенное в соотношениях (9.4) и (9.5). Из-за квантования, при целых числах заполнения $i = 1, 2, 3 \dots$, когда плотность состояний на уровне Ферми обращается в нуль, проводимость σ_{xx}^0 тоже обращается в нуль, а при полужелых $\nu = i + 1/2$ она достигает максимума. Поэтому использовать нижнюю часть полуокружности для определения начальных точек траекторий нельзя. Кривая, которая ее заменяет, условно показана на рис. 9.25, б точками.

На этом можно считать «эскизное проектирование» законченным и приступить к экспериментальной проверке диаграммы.

Заметьте. Поскольку значения n , τ и B определяют начальную точку на потоковой траектории, при их изменении происходит не движение вдоль траектории, а перескок с одной траектории на другую. Поэтому зависимости компонент тензора сопротивления от B или от V_g , самый распространенный тип измерений, с точки зрения проверки структуры диаграммы мало информативны. Например, на рис. 9.21 ясно видно, что фазовый переход, помеченный как $1\downarrow$, сдвинут относительно полуцелого числа заполнения $\nu=3,5$. Однако нетрудно убедиться, что даже эта информация пропадет при переносе экспериментальных точек с рис. 9.21 на фазовую диаграмму.

Для движения вдоль потоковой траектории нужно было бы менять размер системы L , оставаясь при температуре $T = 0$, что конечно практически невозможно. Но в реальном эксперименте при конечной температуре квантовая когерентность теряется на диффузионной длине потери фазы L_φ , определенной формулой (2.7) в гл. 2. Если $L_\varphi < L$, то скейлинговые соотношения определяются именно диффузионной длиной. Это дает принципиальную возможность двигаться вдоль фазовых траекторий за счет изменения температуры. Для этого дополнительно требуется, чтобы на всем температурном интервале диффузионная длина была много больше упругой длины свободного пробега l , так чтобы при любой температуре имели место неравенства

$$l \ll L_\varphi(T) < L. \quad (9.38)$$

Вспомните: В гл. 6 при обсуждении того, как слабая локализация отражается на скейлинговых кривых, описывающих переход металл–изолятор, мы использовали обратную замену: полагая, что температура $T = 0$, мы подставляли в формулы слабой локализации L вместо L_φ .

Начнем с описания эксперимента, сделанного на допированных кремнием пленках GaAs с низкой подвижностью [28], у которых во всех достижимых магнитных полях справедливо неравенство $\Omega\tau < 1$. Пленки GaAs толщиной от 25 до 40 нм с концентрацией около $1,5 \cdot 10^{17} \text{ см}^{-3}$ кремния, игравшего роль донорной примеси, были выращены методом молекулярной эпитаксии. Такой способ приготовления образцов обеспечивает более однородное распределение примесей и электронной концентрации вдоль 2D-слоя. По измеренным при некоторой фиксированной температуре величинам ρ_{xx} и ρ_{xy} вычисляли значения σ_{xx} и σ_{xy} . Их делили каждое на два, предполагая, что электронная система состоит из двух не взаимодействующих подсистем

с разными направлениями спинов, вносящих одинаковый вклад в проводимость, и наносили на фазовую плоскость точку с координатами

$$(\sigma_{xy}^\uparrow = \sigma_{xy}/2, \sigma_{xx}^\uparrow = \sigma_{xx}/2).$$

Серии из одинаковых значков на рис. 9.26 — это результаты измерений одной пленки в фиксированном поле, но при разных температурах. Разные значки означают разные поля от 0,9 Тл до 6 Тл, а разному заполнению значков (значкам, зачерненным полностью, снизу, справа или пустым) соответствуют четыре разные пленки. Эксперимент подтвердил предсказываемую теорией общую структуру и форму траекторий потока в окрестности межфазовой границы $0 \leftrightarrow 1$, в том числе и в области $\Omega\tau < 1$.

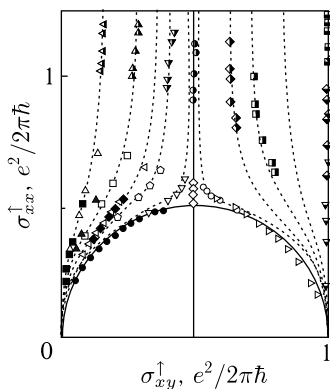


Рис. 9.26. Скейлинговая фазовая диаграмма сильно легированных пленок GaAs [28]. Смысл разных значков пояснен в тексте (серии из одинаковых значков получены путем изменения температуры в интервале от 4,2 К до 40 мК). Пунктирные линии фазовых траекторий и сепаратриса в виде полуокруга $(\sigma_{xy} - 1/2)^2 + \sigma_{xx}^2 = 1/4$ построены на основании скейлинговой теории [22, 25]

Вместе с тем, важнее проверить *принципиальные* предсказания модели двухпараметрического скейлинга для области $\Omega\tau < 1$, и здесь успехи пока скромные. Будем двигаться по этой области диаграммы справа налево, например, двигаться, уменьшая магнитное поле, вдоль верхней части полуокружности $\sigma_0 = \text{const}$ (рис. 9.25, б). При пересечении фазовых границ $\sigma_{xy} = i + 1/2$, $i = \dots, 2, 1, 0$, величина холловской проводимости σ_{xy} должна уменьшаться на $e^2/2\pi\hbar$ так же, как при движении вдоль нижней части она увеличивалась. Это означает, что кривая $\rho_{xy}(B)$ должна подыматься ступенями вверх не только при росте поля в области больших полей, как на рис. 9.7, но и при его уменьшении в области малых полей (см. уравнение (9.34) и рис. 9.22). Такой *обратной* ступенчатой структуры функции $\rho_{xy}(B)$ в малых полях до сих пор никому не удалось наблюдать. В терминах всплывания уровней это означает, что не удалось наблюдать, как протяженные уровни, всплывая, *поочередно, по-одному* пересекают уровень Ферми (ср. рис. 9.22 предыдущего параграфа).

Эксперименты на более чистых образцах, у которых в сильных полях достигается предел $\Omega\tau > 1$, тоже не вносят полной ясности относительно двухпараметрического скейлинга. Результаты одних ра-

бот согласуются со скейлинговой диаграммой, а результаты других — нет. Известен ряд утверждений, что вопреки диаграмме двухпараметрического скейлинга, наблюдаются переходы из состояния $i = 0$ непосредственно в состояния $i = 2$, $i = 3$ и т. д. Однако эти утверждения основываются на предположении, что установить сам факт фазового перехода и идентифицировать фазы можно, исходя из знака производных $\partial \rho_{xx} / \partial T$ при низких температурах. Это предположение само нуждается в доказательствах и проверке.

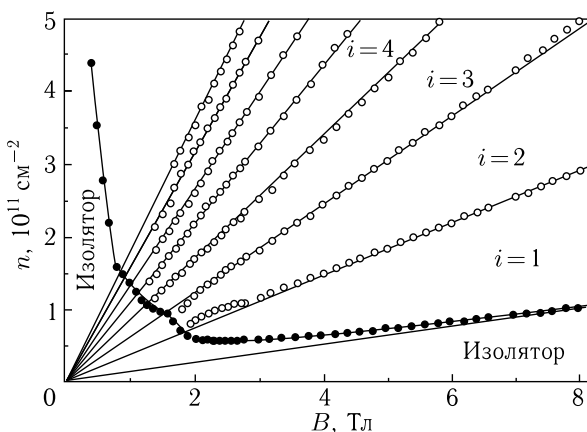


Рис. 9.27. Переходы между различными фазами двумерного электронного газа в гетероструктуре Ge/SiGe в сильном магнитном поле на фоне веера полузаполненных уровней Ландау. Белые точки получены из положения ступенек функции ρ_{xy} , черные — из пересечения изотерм $\rho_{xx}(B)$ при разных температурах, прямые построены в соответствии с уравнением (9.30) [29]

Проблема переходов $0 \rightarrow (i > 1)$ тесно связана с вопросом, всплывает ли только самый нижний из протяженных уровней (что собственно и видно на рис. 9.23 и 9.24) или следующие протяженные уровни всплывают тоже. Это хорошо видно из диаграммы на рис. 9.27, взятой из работы [29]. Здесь белые точки поставлены на основании положения ступенек в ρ_{xy} , которые находятся на переходах между двумя различными квантовыми жидкостями. Черные точки получены из анализа производных $\partial \rho_{xx} / \partial T$. Из сравнения с рис. 9.23 и 9.24 ясно, что это, по существу, есть всплывание нижнего протяженного уровня. Если ветви белых точек упираются в черную ветвь и уровни сливаются, то появляются участки границ между изолятором $i = 0$ и состояниями $i > 1$, и возможны соответствующие переходы $0 \rightarrow (i > 1)$. Если эти ветви загибаются вверх, как на рис. 9.22, то переходы $0 \rightarrow (i > 1)$ невозможны.

Однако, поскольку в любом случае состояние изолятора примыкает к оси ординат, фазовая диаграмма КЭХ согласуется при предельном переходе $B \rightarrow 0$ с утверждением скейлинговой гипотезы о локализации электронного 2D-газа в нулевом магнитном поле.

9.8. Контрольные вопросы и задачи

1. Где течет транспортный ток в режиме квантового эффекта Холла: только по краевым каналам или по всему образцу:

- а) в очень чистом образце сравнительно малого размера?
- б) в более грязном образце большого размера?
- а) в основном по краевым каналам;
- б) по всему образцу (см. рис. 9.15).

2. Вычислить плотность состояний между двумя уровнями Ландау двумерного электронного газа в режиме квантового эффекта Холла по температурным зависимостям продольного сопротивления $R_{xx}(T)$ в разных магнитных полях.

Список литературы

1. Квантовый эффект Холла / Под ред. Р. Е. Пренджа и С. М. Гирвина. — М.: Мир, 1989.
2. Physics of low dimensional structures / Ed. by B. Butcher, N. H. March, M. P. Tosi. — Plenum Press, 1993.
3. Демиковский В. Я., Вугальтер Г. А. Физика квантовых низкоразмерных структур. — М.: Логос, 2000.
4. Абрикосов А. А. Основы теории металлов. — М.: Физматлит, 2006.
5. Гантмахер В. Ф., Левинсон И. Б. Рассеяние носителей тока в металлах и полупроводниках. — М.: Наука, 1984.
6. Klitzing K. v., Dorda G., Pepper M. // Phys. Rev. Lett. **45**, 494 (1980).
7. Краснополин И. Я., Пудалов В. М., Семенчинский С. Г. // ПТЭ №6, 5 (1987)
8. Mohr P. J., Taylor B. N. // Rev. Mod. Phys. **72**, 351 (2000)
9. Ebert G., Klitzing K. v., Probst C., Ploog K. // Solid State Commun. **44**, 95 (1982).
10. Koch S., Haug R. J., Klitzing K. v., Ploog K. // Phys. Rev. B. **43**, 6828 (1991).
11. Ebert G., Klitzing K. v., Probst C., Schuberth E., Ploog K., Weimann G. // Solid State Commun. **45**, 625 (1983).
12. Iordansky S. V. // Solid State Commun. **43**, 1 (1982).
13. Kazarinov R. F., Luryi S. // Phys. Rev. B **25**, 7626 (1982); Phys. Rev. B **27**, 1386 (1983).
14. Büttiker M. // Advances in Solid State Physics, v.**30**, 40 (1990).
15. Chklovskii D. B., Shklovskii B. I., Glazman L. I. // Phys. Rev. B. **46**, 4026 (1992).
16. Thouless D. J. // Phys. Rev. Lett. **71**, 1879 (1993).

17. Долгополов В. Т., Житенев Н. Б., Шашкин А. А. Письма в ЖЭТФ **52**, 826 (1990); *Dolgoplov V. T., Shashkin A. A., Zhitenev N. B., Dorozhkin S. I., Klitzing K. v.* // Phys. Rev. B. **46**, 12560 (1992).
18. *Stahl E., Weiss D., Weimann G., Klitzing K. v., Ploog K.* // J. Phys. C: Solid State Phys. **18**, L783 (1985).
19. *Tsui D. C., Störmer H. L., Gossard A. C.* // Phys. Rev. B. **25**, 1405 (1982).
20. *Kivelson S., Lee D.-H., Zhang S.-C.* // Phys. Rev. B. **46**, 2223 (1992).
21. *Paalanen M. A., Tsui D. C., Gossard A. C.* // Phys. Rev. B. **25**, 5566 (1982).
22. *Pruisken A. M. M.* в книге [1].
23. *Glozman I., Johnson C. E., Jiang H. W.* // Phys. Rev. Lett. **74**, 594 (1995).
24. *Khmelnitskii D. E.* // Phys. Lett. **106A**, 182 (1984); *Helvetica Physica Acta.* **65**, 164 (1992).
25. *Huckestein B.* // Rev. Mod. Phys. **67**, 357 (1995).
26. *Huckestein B.* // Phys. Rev. Lett. **84**, 3141 (2000).
27. *Khmelnitskii D. E.* // JETP Lett. **38**, 552 (1982).
28. *Murzin S. S., Weiss M., Jansen A. G. M., Ebert K.* // Phys. Rev. B. **66**, 233314 (2002).
29. *Hilke M., Shahar D., Song S. H., Tsui D. C., Xie Y. H.* // Phys. Rev. B. **62**, 6940 (2000).

Глава 10

КВАНТОВЫЕ ФАЗОВЫЕ ПЕРЕХОДЫ ¹⁾

Выше мы уже дважды встречались с квантовыми фазовыми переходами: в гл. 5 и 6 речь шла о переходах металл–изолятор, в гл. 9 обсуждались переходы между разными состояниями квантовой холловской жидкости. Принципиальным в обоих случаях является то, что переходы могут происходить при абсолютном нуле температуры $T = 0$ (см. в связи с этим дискуссию в начале гл. 5). Именно по этому признаку их относят к *квантовым* переходам, описывающим изменение *основного состояния*, когда некоторый управляющий параметр x принимает критическое значение x_c .

В принципе существуют два типа квантовых переходов. Точка квантового перехода может быть конечной точкой линии термодинамических переходов на плоскости (x, T) , т.е. на плоскости x — управляющий параметр, T — температура: $x_c(T) \rightarrow x_c(0)$ при $T \rightarrow 0$. Например, обычно так ведут себя магнитные переходы и переходы сверхпроводник-изолятор с магнитным полем в качестве управляющего параметра. Квантовые переходы, которые обсуждаются здесь, в принципе реализуются только при $T = 0$, а при конечной температуре превращаются в плавный кроссовер (см., например, рис. 6.3 в гл. 6). Поэтому они изображаются изолированной точкой x_c на оси абсцисс плоскости (x, T) . Мы ограничимся здесь именно этим случаем.

Мы начнем изложение материала этой главы с обсуждения на качественном уровне тех идей, которые были заложены в основу при построении теории квантовых фазовых переходов. Затем посмотрим, в какой мере в общую теоретическую схему укладываются известные нам данные о переходах металл–изолятор и о переходах между разными квантовыми холловскими жидкостями.

¹⁾ Глава является сокращенным вариантом обзора [1] в УФН. Желаящим освоить математический аппарат теории квантовых фазовых переходов следует обратиться к книге [2]; с основными идеями можно ознакомиться также в обзорах [3, 4]. Аналогичную исходную систему понятий, постановку вопросов, приемы рассмотрения можно найти в книге о термодинамических фазовых переходах [5].

10.1. Параллели и различия между классическим и квантовым фазовыми переходами

Прежде чем обсуждать квантовые фазовые переходы по существу, кратко напомним некоторые положения статистической физики. Будем предполагать, что мы имеем дело с макроскопической системой, состоящей из огромного числа частиц. Рассматриваемая система почти изолирована и во многих отношениях ведет себя как замкнутая, однако у нее имеется возможность обмениваться частицами и энергией с еще большей системой, служащей резервуаром, т. е. наша система является подсистемой резервуара.

Центральное место в описании систем, находящихся в тепловом равновесии, занимает статистическая сумма

$$Z = \sum_i e^{-\varepsilon_i/T}, \quad (10.1)$$

в которой суммирование происходит по всем возможным состояниям системы, а ε_i — энергии этих состояний. Вообще говоря, функция Z аналитическая, но в точках фазовых переходов она теряет аналитичность и имеет особенности.

С помощью функции Z можно вычислять значения физических величин, характеризующих систему в равновесии. В частности, через статистическую сумму Z выражается вероятность $p_i \propto \exp(-\varepsilon_i/T)$ пребывания системы в состоянии с энергией ε_i :

$$p_i = \frac{e^{-\varepsilon_i/T}}{\sum_i e^{-\varepsilon_i/T}} = \frac{e^{-\varepsilon_i/T}}{Z}. \quad (10.2)$$

То обстоятельство, что энергия подсистемы не фиксирована, а подчиняется распределению (10.2), означает наличие классических тепловых флуктуаций.

Для перехода к квантовомеханическому описанию предположим сначала, что подсистема полностью замкнута и никак не взаимодействует с большой системой. Это позволяет воспользоваться стандартным аппаратом квантовой механики и применить к подсистеме стационарное уравнение Шредингера. Получившийся набор стационарных значений энергии ε_i и соответствующий набор волновых функций $\varphi_i(q)$ системы можно считать атрибутами подсистемы в нулевом приближении. В выражении (10.1) фигурируют именно эти энергии ε_i . Набор функций $\varphi_i(q)$ удобен тем, что он имеет отношение к рассматриваемой подсистеме и в то же время является полным, так что по нему удобно производить разложение.

Формулы (10.1) и (10.2) можно переписать с помощью функций $\varphi_i(q)$. Обозначим через q точку в координационном пространстве $\mathcal{Q}$ обобщенных координат нашей системы. Вероятность того, что система

находится в точке q , будучи в этом состоянии, равна $\varphi_i^*(q)\varphi_i(q)$. Если $\varphi_i(q)$ лишь одно из многих возможных состояний системы, то полная вероятность $P(q)$ найти ее в точке q получается суммированием по всем состояниям:

$$P(q) = \frac{1}{Z} \sum_i \varphi_i^*(q)\varphi_i(q)e^{-\varepsilon_i/T}. \quad (10.3)$$

Поскольку интеграл $\int P(q) dq = 1$, статистическая сумма Z равна сумме интегралов

$$Z = \sum_i \int \varphi_i^*(q)\varphi_i(q)e^{-\varepsilon_i/T} dq, \quad (10.4)$$

причем суммирование происходит по полной системе функций $\{\varphi_i\}$. Впрочем, соотношение (10.4) можно было получить и непосредственно из уравнения (10.1), умножив в нем каждый член суммы на тождественно равный единице интеграл $\int \varphi_i^*(q)\varphi_i(q) dq$ и внося множитель $e^{-\varepsilon_i/T}$ под знак интеграла.

Строго говоря, из-за взаимодействия с внешним окружением подсистема не может быть описана волновой функцией. Вместо нее для описания подсистемы вводится матрица плотности

$$\rho(q', q) = \int \psi^*(q', X)\psi(q, X) dX, \quad (10.5)$$

где q соответствует совокупности координат подсистемы, X — оставшиеся координаты системы, а $\psi(q, X)$ — волновая функция замкнутой системы.

Средние значения $\langle s \rangle$ любой физической величины вычисляются теперь не с помощью волновой функции, а с помощью матрицы плотности

$$\langle s \rangle = \frac{1}{A} \int \{\hat{s}[\rho(q', q)]\}_{q'=q} dq, \quad A = \int \rho(q, q) dq. \quad (10.6)$$

В этом выражении сначала следует применить оператор $\hat{s}$, действующий на функции от переменной q , к $\rho(q', q)$, затем положить $q' = q$ и проинтегрировать. Оператор, который формально входит в выражение для нормировочной константы A , тождественно равен единице; поэтому в подынтегральном выражении сразу учтено, что $q' = q$.

В соответствии с общим правилом (10.6), среднее значение координаты $\langle q \rangle$, например, равно

$$\langle q \rangle = \frac{1}{A} \int \int \psi^*(q, X)q\psi(q, X) dq dX = \frac{\int q\rho(q, q) dq}{\int \rho(q, q) dq}. \quad (10.7)$$

Физический смысл матрицы плотности можно прояснить, выписав ее в явном виде для подсистемы, находящейся в статистическом равновесии при конечной температуре T :

$$\rho(q', q) = \sum_i \varphi_i^*(q') \varphi_i(q) e^{-\varepsilon_i/T}. \quad (10.8)$$

Из формул (10.7) и (10.8) получим для подсистемы в статистическом равновесии

$$\begin{aligned} \langle q \rangle &= \frac{1}{A} \sum_i \int \varphi_i^*(q) \varphi_i(q) e^{-\varepsilon_i/T} q dq = \frac{1}{A} \sum_i \langle q \rangle_i e^{-\varepsilon_i/T}, \\ \langle q \rangle_i &= \int \varphi_i^*(q) \varphi_i(q) q dq. \end{aligned} \quad (10.9)$$

Проведенное в (10.9) усреднение с помощью матрицы плотности включает в себя как результат вероятностного описания в квантовой механике в виде $\langle q \rangle_i$, так и неполную информацию о системе (статистическое усреднение).

Пользуясь функциями $\varphi_i(q)$, функцию $\rho(q', q)$ можно записать в матричном виде

$$\rho(q', q) = \|\rho_{ij}\| = \left\| \int \varphi_j^*(q) \rho(q', q) \varphi_i(q) dq \right\|. \quad (10.10)$$

Нормированная матрица плотности

$$w_{ij} = \frac{\rho_{ij}}{\sum_i \rho_{ii}} \quad (10.11)$$

называется статистической матрицей.

Подставив с уравнение (10.6) вместо $\hat{s}$ оператор энергии $\hat{H}$, получим

$$\langle \varepsilon \rangle = \sum_i w_{ii} \varepsilon_i. \quad (10.12)$$

Это означает, что вероятность p_i обнаружить у нашей подсистемы энергию ε_i равна диагональному элементу матрицы w_{ii}

$$p_i = w_{ii}. \quad (10.13)$$

Выражение (10.13) является квантовым аналогом соотношения (10.2).

Зависимость волновой функции подсистемы не только от ее внутренних обобщенных координат q_{in} , но и от внешних по отношению к ней координат системы q_{out} сохраняется и при $T = 0$. Поэтому даже при $T = 0$ у рассматриваемой подсистемы нет четко зафиксированного основного состояния, в котором она должна все время находиться. Именно это позволяет реализовываться *квантовым флуктуациям*.

При этом измеряемая энергия подсистемы тоже может флуктуировать в рамках, допускаемых соотношением неопределенности

$$\delta\varepsilon\tau \sim \hbar. \quad (10.14)$$

В частности, наличие квантовых флуктуаций означает, что при $T = 0$ в одной из фаз вблизи точки перехода в течение времени τ могут существовать области чужой фазы с характерными размерами ξ . Как обычно, корреляционная длина ξ зависит от расстояния до перехода $\delta x = |x - x_c|$, а на самом переходе расходится. Энергия $\delta\varepsilon$, необходимая для появления такой области, зависит от расстояния до перехода δx и от размера ξ . Она в свою очередь ограничивает время τ : время τ тем меньше, чем больше энергия квантовой флуктуации.

На математическом языке сказанное означает, что, в отличие от классического выражения (10.2), квантовое выражение (10.13) для p_i не стремится экспоненциально к нулю по мере понижения температуры, а приближается к константе, зависящей не от температуры, а от ε_i , размеров и свойств подсистемы (см. рис. 10.1, а).

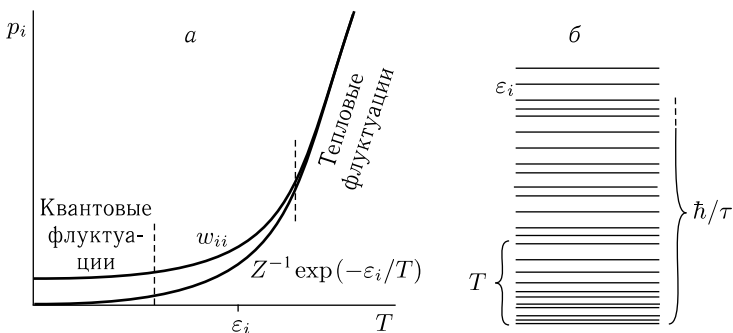


Рис. 10.1. а) Низкотемпературное поведение диагональных элементов w_{ii} статистической матрицы, соответствующих малым значениям энергии; б) одновременное возбуждение тепловых и квантовых флуктуаций

Относительную роль тепловых и квантовых флуктуаций поясняет рис. 10.1 б. На нем изображен спектр возбуждений подсистемы в одной из фаз, т. е. вдали от точки $x = x_c$ фазового перехода. Скобка слева показывает масштаб, связанный с температурой. Классически возбужденными являются только состояния с частотами $\omega_i \lesssim T/\hbar$. Диапазон тепловых состояний ограничен сверху температурой, но зато у этих состояний большие числа заполнений. Состояния $\omega_i > T/\hbar$ возбуждаются в основном за счет квантовых процессов и имеют малые числа заполнений. Однако, чем ниже температура, тем роль квантовых возбуждений больше.

Классические статистические суммы в виде (10.1) или (10.4) содержат только члены, описывающие тепловые флуктуации, каждый из

которых имеет множитель $\exp(-\varepsilon/T)$. Для теоретического описания квантового фазового перехода, аналогичного теории термодинамического перехода, нужно построить выражение, которое бы играло роль статистической суммы, но содержало бы члены, описывающие квантовые флуктуации. Построением такого выражения мы сейчас и займемся.

Статистическую сумму Z , записанную в виде (10.4), можно считать суммой диагональных матричных элементов оператора $\exp(-\hat{H}/T)$ на полном наборе собственных функций φ_i . Этот оператор очень похож на оператор

$$\hat{S} = \exp\left(-\frac{i}{\hbar}\hat{H}t\right), \quad (10.15)$$

который описывает эволюцию квантовой системы во времени. Собственную функцию гамильтониана $\hat{H}$ этот оператор лишь умножает на фазовый множитель:

$$\varphi_i(t) = \hat{S}\varphi_i(0) = \exp\left(-\frac{i}{\hbar}\hat{H}t\right)\varphi_i(0) = \exp\left(-\frac{i}{\hbar}\varepsilon_i t\right)\varphi_i(0),$$

но на произвольную функцию $\Psi(q, 0) = \sum_i a_i \varphi_i(q)$ он воздействует более радикально:

$$\Psi(q, t) = \hat{S}\Psi(q, 0) = \hat{S} \sum_i a_i \varphi_i(q) = \sum_i a_i e^{-i\varepsilon_i t/\hbar} \varphi_i(q), \quad (10.16)$$

поскольку члены суммы (10.16) приобретают разные изменения фазы.

У оператора $\hat{S}$ диагональные матричные элементы

$$\langle \Psi_i | \exp(-i\hat{H}t/\hbar) | \Psi_i \rangle$$

очень похожи на матричные элементы (10.4). Совпадение станет полным, если сделать замену переменных:

$$it/\hbar = 1/T. \quad (10.17)$$

Подстановка (10.17) позволяет по-новому интерпретировать матричные элементы (10.4). Квадрат элемента $\langle i | \exp(-\hat{H}/T) | i \rangle$ можно рассматривать как вероятность того, что подсистема, стартовав из состояния $\langle i |$ и эволюционируя под действием оператора $\hat{S}$, вернется в исходное состояние $|i\rangle$ через «мнимое время» $\hbar/iT$. При этом следует иметь в виду, что вследствие открытости подсистемы ее волновая функция не является стационарной, $|i\rangle \equiv \Psi(q, 0) = \Psi(q, \hbar/T) \neq \varphi_i(q)$; это относится и к промежуточным состояниям $|j\rangle \equiv \Psi(q, t)$, где $t < \hbar/T$.

Чтобы отличать мнимое время от реального, введем для него обозначение $\tilde{t}$. Мнимое время $\tilde{t}$ является параметром, который управляет квантовой эволюцией волновой функции подсистемы. Оно изменяется в ограниченном диапазоне

$$0 < |\tilde{t}| < \hbar/T. \quad (10.18)$$

Допустимый интервал мнимого времени тем больше, чем ниже температура, что в соответствии с рис. 10.1 означает, что по мере понижения температуры квантовые флуктуации играют роль при все более низких энергиях.

Пусть подсистема побывала за время $i\hbar/T$ в какой-то последовательности из N состояний, оставаясь в каждом время $\delta_k \tilde{t}$, и в конце, при $\sum_{k=0}^N \delta_k \tilde{t} = i\hbar/T$, вернулась в состояние $\langle i|$. Такая замкнутая траектория в пространстве состояний описывается произведением матричных элементов

$$P_{i\alpha} = \langle i| \exp(-i\hat{H}\delta_0\tilde{t}/\hbar)|j_1\rangle \langle j_1| \exp(-i\hat{H}\delta_1\tilde{t}/\hbar)|j_2\rangle \dots \dots \langle j_N| \exp(-i\hat{H}\delta_N\tilde{t}/\hbar)|i\rangle, \quad (10.19)$$

а на рис. 10.2 изображается линией, концы которой располагаются на противоположных сторонах полосы шириной $\hbar/T$ точно друг под другом. Из рисунка ясно, что таких траекторий может быть очень много и

$$\langle i| \exp(-\hat{H}/T)|i\rangle = \sum_{\alpha} P_{i\alpha}. \quad (10.20)$$

После подстановки в описание (10.19) любой из них вместо функций $\langle j|$ их разложений (10.16) по полной системе функций $\{\varphi_i\}$ и интегрирования, в выражении останутся, вследствие ортогональности функций φ_i , только диагональные элементы типа $\langle \varphi_i| \exp(-i\hat{H}\delta_0\tilde{t}/\hbar)|\varphi_i\rangle$, которые как раз и должны входить в статистическую сумму Z .

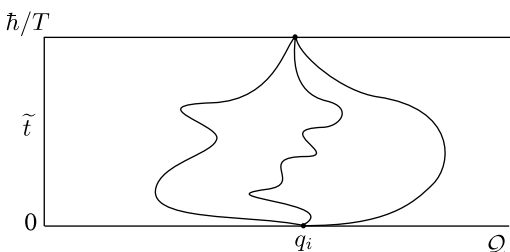


Рис. 10.2. Различные траектории подсистемы в координационном пространстве $\{Q\tilde{t}\}$ в интервале изменения переменной $|\tilde{t}|$ от 0 до $\hbar/T$. Предполагается, что вдоль оси абсцисс располагаются все возможные состояния реальной подсистемы в пространстве размерности d

Произведение матричных элементов $P_{i\alpha}$ соответствует цепочке последовательных виртуальных переходов. Суммирование по α означает, что учтены все возможные замкнутые цепочки. Произведенная при помощи формулы (10.20) модернизация выражения для Z означает, что мы учли квантовые свойства системы, добавив к реальным переходам виртуальные. Однако, множество значений энергии $\{\varepsilon\}$, по

которому происходит суммирование, теперь другое, потому что энергия меняется теперь не только при изменении координат q , но и со временем $\hat{t}$. Такого числа разных состояний в нашей реальной системе нет. Однако, построим воображаемую классическую систему с большей размерностью с тем, чтобы показать, что *поведение квантовой системы в окрестности точки перехода при конечной температуре в пространстве размерности d аналогично поведению классической системы в пространстве большей размерности $\mathcal{D} > d$* .

Путь к такому построению указывает рис. 10.2. Добавим к d осям исходного пространства дополнительную ось, вдоль которой отложим мнимое время $\hat{t} = it$. В графических образах рис. 10.2 это означает, что к оси абсцисс добавлена ось ординат. Полоса (10.18) на рис. 10.2, в которой находится изображающая точка, тем шире, чем ниже температура. При $T = 0$ полоса превращается в полуплоскость, а рост температуры, наоборот, сужает полосу и уменьшает ее вклад в статистические свойства квантовой системы.

Конечно, в общем виде получившаяся математическая конструкция абсолютно непрактична. Однако, как это обычно бывает при обосновании скейлинговых соотношений, для нас важно не вычисление конкретной статистической суммы, а некоторые ее принципиальные свойства в окрестности точки перехода. Как мы увидим, таким свойством является анизотропия $\mathcal{D}$ -мерного пространства, т. е. неравноценность его осей.

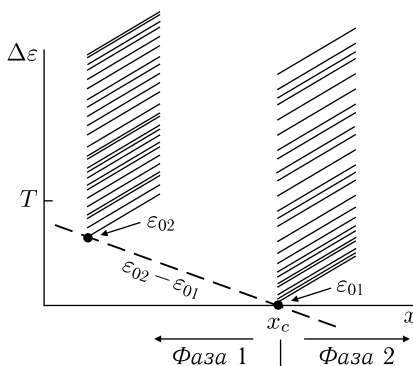


Рис. 10.3. Зависимость разности наименьших возможных значений энергии двух фаз $\Delta \varepsilon = \varepsilon_{02} - \varepsilon_{01}$ от управляющего параметра (штриховая прямая). При значении управляющего параметра $x = x_c - \delta x$ происходит флуктуационное возбуждение фазы 2 внутри равновесной фазы 1. На рисунке изображена ситуация, когда основную роль в этом возбуждении играют тепловые флуктуации, благодаря тому, что $|\Delta \varepsilon| < T$

В точке термодинамического фазового перехода статистическая сумма Z имеет особенность. Чувствительность функции (10.4) к наличию фазового перехода обусловлена тем, что при приближении

к переходу в энергетический интервал T , определяющий существенные члены суммы (10.4), попадают не только уровни φ_i из набора, соответствующего равновесной фазе, но и уровни φ'_i неравновесной фазы (см. рис. 10.3). Это обуславливает возможность флуктуационных переходов между уровнями $|i\rangle$ и $|i'\rangle$ из разных наборов. После введения квантовых флуктуаций в статистическую сумму появились дополнительные возможности. Если температура низкая, так что $|\Delta\varepsilon| > T$, то флуктуационное появление другой фазы все равно возможно, но только за счет квантовых флуктуаций.

10.2. Критическая окрестность квантового перехода

Вблизи классического фазового перехода всегда существует окрестность на фазовой плоскости, в которой все физические величины зависят только от корреляционной длины ξ , которая представляет собой характерный размер тепловых флуктуаций. Эту окрестность называют критической или скейлинговой областью. Вблизи квантового фазового перехода, *при* $T = 0$, тоже существует область значений управляющего параметра δx , в которой физические величины выражаются через длину ξ , пропорциональную некоторой степени $|\delta x|$. Эта величина определяет характерный размер квантовых флуктуаций. При конечной температуре $T \neq 0$, однако, скейлинговая область квантового фазового перехода выглядит сложнее.

Пространство $\{^d X, \tilde{t}\}$ имеет размерность $\mathcal{D}$ из-за дополнительного подпространства мнимого времени. Введем в нем корреляционные длины. Сохранив традиционное обозначение ξ для корреляционной длины в обычном d -мерном подпространстве, корреляционную длину вдоль дополнительных осей обозначим ξ_φ . Используемый индекс призван напоминать, что величина ξ_φ связана с квантовым аспектом задачи и со спецификой волновых функций.

Заметьте. Величина ξ_φ имеет размерность не длины, а $\hbar/T$, т. е. измеряется не в сантиметрах, а в секундах. Однако поскольку подпространство мнимого времени входит в $\mathcal{D}$ -мерное пространство, то «длина» в соответствующих направлениях выступает в той же роли, что и настоящая длина ξ .

Обе корреляционные длины при приближении к переходу, т. е. при $x \rightarrow x_c$ и $T \rightarrow 0$, будут расходиться. Согласно теории термодинамических непрерывных переходов, расходимость описывается степенными функциями, но показатели степени у двух корреляционных длин могут быть, вообще говоря, разные. Обычно это записывается в таком виде:

$$\xi \propto \delta x^{-\nu}, \quad \delta x = |x - x_c|, \quad (10.21)$$

$$\xi_\varphi \propto \xi^z. \quad (10.22)$$

Показатели степени ν и z называются критическими индексами, причем к названию индекса z добавляют слово "динамический". И названия, и обозначения происходят из теории термодинамических переходов. В частности, динамический критический индекс входит там в соотношение, связывающее время жизни тепловых флуктуаций с их размером: $\tau \propto \xi^z$. В квантовой задаче при $T = 0$ все выглядит очень похоже: длина ξ характеризует пространственные корреляции, т. е. характерный размер квантовых флуктуаций, а время ξ_φ — временные корреляции.

Формула (10.22) сразу задает размерность пространства, в котором должна находиться воображаемая классическая система. Вдоль всех осей этого пространства должна быть отложена настоящая длина, т. е. из оси на рис. 10.2, вдоль которой отложено $\tilde{t}$, нужно сделать оси, вдоль которых отложена длина. По соображениям размерности из соотношения (10.22) следует, что длина, эквивалентная корреляционной псевдо-длине ξ_φ , пропорциональна

$$L_\varphi \propto \xi_\varphi^{1/z}, \quad (10.23)$$

так что элемент объема в интересующем нас пространстве пропорционален

$$(d\xi)^d d\xi_\varphi \propto (d\xi)^d (dL_\varphi)^z, \quad (10.24)$$

и подпространство мнимого времени получается z -мерным, но с обычными «пространственными» координатами. Отсюда

$$\mathcal{D} = d + z. \quad (10.25)$$

Квантовый фазовый переход в пространстве размерности d эквивалентен классическому фазовому переходу в пространстве размерности $\mathcal{D}$.

Следствия уравнений (10.21) и (10.22) удобно обсуждать, используя плоскость (x, T) — см. рис. 10.4. При $T = 0$, на оси абсцисс плоскости (x, T) , управляющий параметр x влияет только на одну независимую корреляционную длину, ξ ; величину ξ_φ (и длину L_φ) можно формально получить из ξ при помощи уравнения (10.22). На некотором отрезке $[x_1, x_2]$, включающем в себя точку x_c , все физические величины выражаются только через ξ .

Пусть теперь температура $T \neq 0$. Будем двигаться по плоскости (x, T) вдоль горизонтальной линии $T \neq 0$ (линия aa на рис. 10.4) в направлении $x \rightarrow x_c$. При некотором значении δx величина ξ_φ , изменяющаяся в соответствии с соотношениями (10.21), (10.22), достигает своего

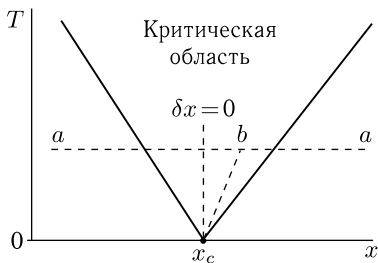


Рис. 10.4. Форма критической области квантового фазового перехода, изображаемого изолированной точкой на плоскости (x, T)

максимально возможного значения $\hbar/T$, обусловленного неравенством (10.18). При меньших значениях δx на линии aa соотношение (10.22) уже не выполняется, потому что ξ растет в соответствии с соотношением (10.21), а ξ_φ остается равным своему максимально возможному значению $\hbar/T$. Параметры ξ и ξ_φ становятся *взаимно независимыми*. Область, в которой имеет место эта независимость, называется критической. Если теперь двигаться к переходу, оставаясь внутри критической области (например, вдоль линии bx_c), то на переходе по-прежнему расходятся оба параметра, но расходимость одного контролируется величиной δx : $\xi \propto \delta x^{-\nu}$, а расходимость другого контролируется температурой: $\xi_\varphi = \hbar/T$.

При низких температурах по обе стороны от критической области находятся скейлинговые области, где в перемешивании состояний двух фаз существенную роль играют квантовые флуктуации (как и непосредственно на оси x при $T = 0$). Как видно из рис. 10.3, это происходит благодаря условию

$$|\varepsilon_02 - \varepsilon_01| > T. \quad (10.26)$$

При понижении температуры уменьшается интервал значений управляющего параметра δx , в котором это условие нарушается. Отсюда необычная форма критической области, ее расширение за счет скейлинговых областей при росте температуры.

Осмыслить и принять эту необычную форму может помочь такое рассуждение. При повышении температуры энергия каждой из фаз вдалеке от точки перехода x_c обычно представляется в виде суммы энергий основного состояния и набора квазичастиц. Естественно, что квазичастицы в разных фазах различны. На некотором отрезке прямой aa на рис. 10.4 над точкой x_c ни то, ни другое описание не годятся. Чем выше температура T , тем дальше от точки перехода проходит прямая и тем длиннее этот отрезок. Соединив границы всех отрезков, получим область на плоскости (x, T) , ограниченную двумя кривыми, исходящими из точки x_c . Вне этой области справедливо одно из двух микроскопических описаний, но внутри ни одно из них неприменимо. Поэтому для нее приходится конструировать феноменологическое скейлинговое описание.

Исходя из формулы (10.23), длину L_φ внутри критической области можно написать в виде

$$L_\varphi \propto (\hbar/T)^{1/z}. \quad (10.27)$$

Для электронных систем это диффузионная длина (2.7), с которой мы сталкивались в гл. 2 и 6. Конечная температура привносит в квантовую задачу характерную энергию T , которая, согласно рис. 10.1, разделяет классические и квантовые флуктуации. Квантовыми являются флуктуации с энергиями $\hbar\omega_\varphi > T$. Их пространственный размер $l_\varphi \propto 1/\omega_\varphi$ ограничен сверху, так как частота $\omega_\varphi > T/\hbar$ ограничена снизу. Дли-

на L_φ как раз и является верхней границей для размера квантовых флуктуаций. Поэтому L_φ часто называют также длиной расфазировки, т. е. длиной, начиная с которой нарушается когерентность в системе электронов.

Появление двух независимых параметров в критической окрестности квантового фазового перехода обусловлено неравенством (10.18). Поэтому скейлинговое описание в критической окрестности квантового фазового перехода называется «скейлинг в слое конечной толщины» («finite-size scaling»). Имеется в виду, что воображаемая термодинамическая система существует в гиперполосе в $\mathcal{D}$ -мерном пространстве с d переменными, меняющимися от 0 до ∞ , и z переменными, меняющимися в интервале от нуля до $(\hbar/T)^{1/z}$.

Внутри критической области $L_\varphi < \xi$, а на границе

$$L_\varphi = \xi. \quad (10.28)$$

Длины L_φ и ξ определены уравнениями (10.21) и (10.22) лишь с точностью до постоянного множителя, но уравнения жестко фиксируют степенную связь между этими переменными. Поэтому уравнение границ критической области имеет вид

$$T = C(\delta x)^{\nu z}, \quad (10.29)$$

где константа C может быть, вообще говоря, разной с двух сторон от перехода.

Поскольку в критической области два независимых параметра, скейлинговые выражения для физических величин усложняются. Мы приведем и прокомментируем только выражения для удельной проводимости и сопротивления

$$\sigma = \xi^{2-d} F(L_\varphi/\xi), \quad \rho = \xi^{d-2} F_1(L_\varphi/\xi), \quad F_1(u) = 1/F(u), \quad (10.30)$$

где $F(u)$ — произвольная функция. Показатель степени первого сомножителя в выражении (10.30) определяется тем, как длина входит в выражения для проводимости при разных размерностях d . Как видно из рис. 10.4, этот сомножитель определяется x -компонентой расстояния до перехода. Отношение L_φ/ξ безразмерное. Именно поэтому правила скейлинга позволяют сделать его аргументом произвольной функции. Отношение зависит и от T , т. е. от y -компоненты расстояния на фазовой плоскости от изображающей точки до перехода, и от расстояния от этой точки до границы критической области вдоль оси x .

В качестве аргумента можно также взять любую степень отношения L_φ/ξ . Это позволяет записать аргумент в разных формах, например:

$$\frac{L_\varphi}{\xi}, \quad \frac{\hbar/T}{\xi^z}, \quad \frac{\hbar/T}{\xi^z}, \quad \frac{\hbar(\delta x)^{z\nu}}{T} \quad \text{или} \quad \frac{\delta x}{(T/\hbar)^{1/z\nu}}. \quad (10.31)$$

Разные формы записи проясняют разные аспекты физического смысла этого отношения. В двух последних формах записи аргумент является

размерным и при этом размерность управляющего параметра заранее не задана; эти формы показывают, как температура входит в аргумент скейлинговой функции.

Хотя физические свойства системы в критической области определяются двумя независимыми параметрами, построенная схема называется однопараметрическим скейлингом, потому что сама система при $T = 0$ полностью определена одним параметром. Для невзаимодействующих электронов это, например, кондактанс системы, который, в свою очередь, задается электронной концентрацией и беспорядком. Однако, уже в сильном магнитном поле формально параметров становится два: поперечный кондактанс и поле, так что скейлинг становится двухпараметрическим.

Двухпараметрическим скейлинг становится и при учете межэлектронного взаимодействия, зависящего от температуры и характерного размера системы. Мы в этой книге стараемся избегать систем взаимодействующих электронов, но один пример системы с таким перенормируемым взаимодействием у нас все же был: спин-орбитальное взаимодействие в параграфе 6.7. Оно эффективно включается лишь когда размер образца L (или диффузионная длина L_φ) становится больше величины $l(\tau_{so}/\tau)^{1/2}$ (см. обозначения в параграфе 6.7 и формулу (2.4)).

Описание схемы применения двухпараметрического скейлинга можно найти в обзоре [6]. Здесь же мы только отметим одно неявное предположение, позволившее построить схему однопараметрического скейлинга: предполагается, что уравнение (10.21) для ξ справедливо не только при $T = 0$, но и при конечных температурах. Благодаря этому происходит разделение влияния переменных:

$$\xi = \xi(x), \quad L_\varphi = L_\varphi(T); \quad (10.32)$$

корреляционная длина ξ зависит только от x , а длина сбоя фазы L_φ только от T .

10.3. Квантовые переходы металл–изолятор

Трехмерный электронный газ. В параграфе 6.3 была построена фазовая диаграмма окрестности 3D-перехода металл–изолятор. Посмотрим теперь на нее с точки зрения понятий и определений, фигурирующих в теории квантовых переходов, см. рис. 10.5. Управляющий параметр x на рис. 6.2 не был конкретизирован, но предполагалось, что имеет место соотношение

$$\xi \propto (\delta x)^{-1}, \quad \text{т. е.} \quad \nu = 1. \quad (10.33)$$

Это предположение сохраняется и на рис. 10.5. Оно не ограничивает универсальность диаграммы, потому что и в формулы (6.17), (6.22) для проводимости, и уравнения (6.18), определяющие линии кроссоверов,

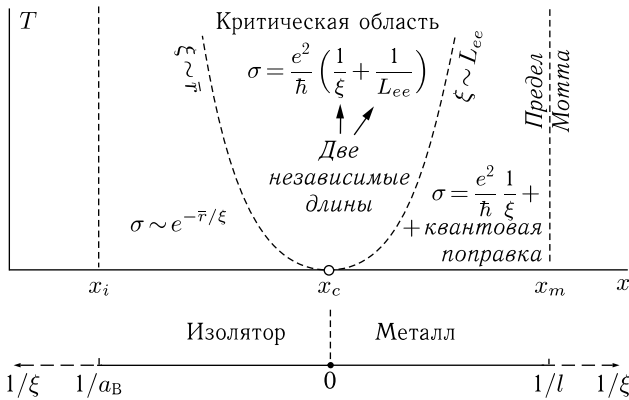


Рис. 10.5. Окрестность квантового фазового перехода трехмерного электронно-газа на плоскости $[x, T]$

все равно входит не x , а ξ . Численные коэффициенты, связывающие ξ и $(\delta x)^{-1}$, определяются тем, что $\xi(x_m) = k_F^{-1}$ и $\xi(x_i) = a_B$.

К точке фазового перехода $(x_c, 0)$ примыкают критическая область, которая расширяется по мере увеличения температуры, как это и было постулировано на рис. 10.4, и две скейлинговые области над отрезками $[x_i, x_c]$ и $[x_c, x_m]$, отделенные от критической области двумя линиями кроссоверов, выходящими из точки $(x_c, 0)$. Из сравнения уравнений этих линий (6.18) и соотношения (10.29) получаем, что динамический критический индекс

$$z = 3. \quad (10.34)$$

Заметьте. Значение (10.34) для динамического критического индекса $z = 3$ получилось благодаря предположению, что расфазировка происходит за счет внутренних процессов в электронном газе без привлечения внешних воздействий, таких как, например, электрон-фононное взаимодействие. (См. аналогичное замечание в гл. 6 после формулы (6.18), а также обсуждение физического смысла величины L_{φ} в обзоре [3].)

Внутри критической области, как и требуется, транспортные характеристики выражаются через два независимых параметра размерности длины: L_{ee} и ξ . В скейлинговых областях параметр только один, а температура существенной роли не играет. Длина расфазировки L_{ee} была введена при анализе проводимости в металлической области (см. формулу (2.30)). Формально она существует и в правой скейлинговой области, но там она контролирует лишь малую квантовую поправку. Собственно говоря, правая граница критической области определяется условием, что эта поправка перестает быть малой (величина L_{ee} на границе критической области предполагается непрерывной). На левой

границе критической области величина $L_\varphi \equiv L_{ee}$ сравнивается со средней длиной прыжка $\bar{r}$.

Выражение (6.15) для проводимости в критической области можно записать в виде, удовлетворяющем общему соотношению (10.30), положив $F(u) = (e^2/\hbar)(1 + 1/u)$:

$$\sigma = \xi^{-1}(e^2/\hbar)(1 + \xi/L_\varphi), \quad (10.35)$$

а формула (6.17) для $\sigma(T)$ получается из (10.35) подстановкой соотношения (10.27). Таким образом фазовую диаграмму рис. 6.2 окрестности 3D-перехода металл–изолятор, построенную на основании скейлинговой гипотезы, исходявшей только из беспорядка и предполагавшей отсутствие взаимодействия, можно полностью описать в терминах теории квантовых переходов.

Перейдем теперь к 2D-системам.

Окрестность виртуального квантового фазового перехода на фазовой плоскости двумерного электронного газа. Согласно теории Абрагамса и др., обсуждавшейся в гл. 6, при отсутствии взаимодействия между электронами в двумерных системах вообще не должно быть перехода металл–изолятор. Однако, линия кроссоверов на плоскости (беспорядок–температура) есть. Действительно, перерисуем рис. 6.8 из гл. 6, направив ось x не вправо, а влево — рис. 10.6. Так

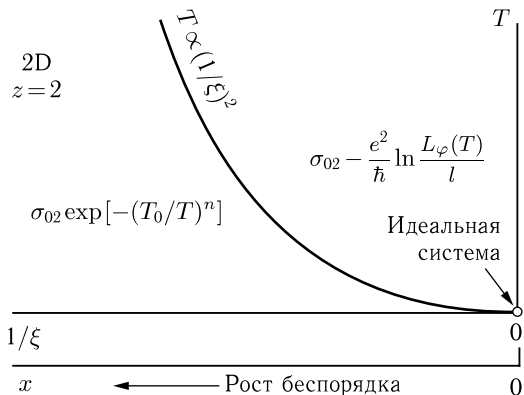


Рис. 10.6. Фазовая плоскость $(1/\xi, T)$ для невзаимодействующего двумерного электронного газа

получившуюся диаграмму будет удобнее сравнивать с диаграммой для 3D-систем на рис. 10.5. Управляющим параметром здесь является некая характеристика беспорядка, такая, что $x = 0$ соответствует его полному отсутствию. Строго говоря, вдоль горизонтальной оси отложена величина $1/\xi$, где ξ — корреляционная длина, определенная формулой (6.33). Поэтому можно сказать, что мы выбрали управляющий параметр таким образом, чтобы выполнялось соотношение (10.33). На диа-

грамме построена кривая (6.34), отделяющая область логарифмических поправок к проводимости от области, где происходит экспоненциальный рост сопротивления. Уравнение этой кривой

$$T = D\hbar/\xi^2, \quad \text{т. е.} \quad T \propto \xi^{-2} \quad (10.36)$$

(здесь D — коэффициент диффузии).

Нетрудно видеть, что рис. 10.6 представляет из себя левую часть фазовой диаграммы окрестности перехода металл-изолятор в трехмерном материале (см. рис. 10.5) с точкой фазового перехода, находящейся на краю диаграммы, в начале координат $T = x = 0$. Для системы невзаимодействующих электронов, локализующихся при сколь угодно малом беспорядке, это естественно: в строго идеальной системе нет оснований для локализации электронов. Когда в гл. 6 говорилось об отсутствии перехода в 2D-системе, подразумевалось, что идеальная система, в которой беспорядок полностью бы отсутствовал, недостижима.

Такая интерпретация графика на рис. 10.6 означает, что область над параболой $T \propto (1/\xi)^2$ является критической областью окрестности квантового перехода. В 2D-системах в формуле (10.30) для проводимости в критической области остается одна скейлинговая переменная u , равная отношению двух характерных длин:

$$\sigma = F(u) \equiv F(L_\varphi/\xi). \quad (10.37)$$

Между тем, в этой области мы обычно пользуемся выражением для проводимости в виде разности между классической частью и квантовой поправкой

$$\sigma = \frac{ne^2 l}{\hbar k_F} - \Delta\sigma = \frac{e^2}{\hbar} \left(a_1 k_F l - a_2 \ln \frac{L_\varphi}{l} \right). \quad (10.38)$$

В формуле (10.38), в отличие от аналогичной формулы (6.32), сохранены численные коэффициенты a_1 и a_2 ; они не выписаны в явном виде, поскольку, вообще говоря, могут зависеть от конкретной модели электронного спектра. Наша цель здесь показать, что наличие этих коэффициентов не мешает представить проводимость (10.38) в виде (10.37).

Действительно, приравняв нулю выражение в скобках в (10.38), получим выражение для ξ (ср. с выводом формулы (6.33)):

$$\xi = l \exp(ak_F l), \quad a = a_1/a_2. \quad (10.39)$$

Подставив в аргумент логарифма в (10.38) вместо l выражение из (10.39), $l = \xi \exp(-ak_F l)$, и вынеся экспоненту из-под знака логарифма, получим

$$\sigma = a_2^{-1} \frac{e^2}{\hbar} \ln \frac{L_\varphi}{\xi}. \quad (10.40)$$

Проводимость σ , включающая в себя классическую часть $\sigma_2 = ne^2 l / \hbar k_F$, зависит только от скейлинговой переменной, как это и должно быть в критической области.

Поскольку уравнение (10.36) описывает границу критической области, из него получается динамический критический индекс $z = 2$. Вспомнив соответствующее значение (10.34) для трехмерных систем, увидим, что для переходов металл–изолятор динамический критический индекс равен размерности, $z = d$. Экспериментальным подтверждением этого соотношения для трехмерных систем являются кривые, представленные на рис. 6.4 и 6.5.

10.4. Квантовые переходы между разными состояниями холловской жидкости ¹⁾

Применим понятия теории квантовых переходов к переходам в системах, находящимся в режиме квантового эффекта Холла, и построим их скейлинговое описание. Поскольку речь пойдет исключительно о двумерных электронных системах, то и сопротивление, и удельное сопротивление, как продольное, так и холловское, все имеют одинаковую размерность [Ом], а соответствующие кондактанс и проводимость — размерность [Ом⁻¹]. Формально поведение системы электронов в сильном магнитном поле при $T = 0$ в окрестности фазовых переходов описывается двумя параметрами, но потоковая диаграмма на рис. 9.25 (раздел 9.7) содержит элемент симметрии: вертикальность сепаратрис, позволяющий упростить рассмотрение. Пусть изображающая точка движется вдоль сепаратрисы вследствие роста L или L_φ , приближаясь к точке квантового перехода. При этом меняется только кондактанс σ_{xx} , а холловская проводимость σ_{xy} остается постоянной. Это позволяет воспользоваться в окрестности перехода формулами однопараметрического скейлинга типа (10.30):

$$R_{uv} = F_{uv}(L_\varphi/\xi), \quad (10.41)$$

где под индексами u и v подразумеваются координаты x или y , а F_{uv} — какие-то неизвестные функции.

Согласно общепринятой модели КЭХ, вблизи центра каждого уровня Ландау существует одно значение энергии ε_c , которому соответствуют делокализованные волновые функции электронов. Переходы между разными состояниями холловской жидкости при $T = 0$ происходят при выполнении условия $\varepsilon_F = \varepsilon_c$. Управляющим параметром x является обычно магнитное поле B или концентрация двумерных электронов n , зависящая от напряжения на затворе V_g :

$$\delta x \equiv |B_c - B| \propto (\delta\varepsilon)^\nu \quad \text{или} \quad \delta x \equiv |n_c - n| \propto |V_{gc} - V_g| \propto (\delta\varepsilon)^\nu. \quad (10.42)$$

¹⁾ Совокупность проблем применения скейлинговых теорий к целочисленному квантовому эффекту Холла подробно обсуждается с теоретической точки зрения в обзоре [7].

Степень ν в обоих случаях одна и та же. Это непосредственно следует из рис. 9.11 и 9.12 и подтверждается тем, что экспериментальные кривые $R_{uv}(B)$ и $R_{uv}(V_g)$, записанные на одном образце в эквивалентных условиях, различаются только масштабом по оси абсцисс (см. гл. 9). Воспользовавшись формулой (10.31), можно написать

$$R_{uv} = F_{uv} \left(\frac{\delta x}{T^{1/z\nu}} \right). \quad (10.43)$$

В отличие от случая переходов металл–изолятор, мы не будем здесь вычислять или предсказывать значения ν и z , а посмотрим, что можно извлечь из эксперимента. На рис. 10.7 приведены в качестве примера записи зависимостей от магнитного поля продольного R_{xx} и поперечного R_{xy} сопротивлений холловского мостика в гетероструктуре на базе GaAs [8], полученные при разных температурах.

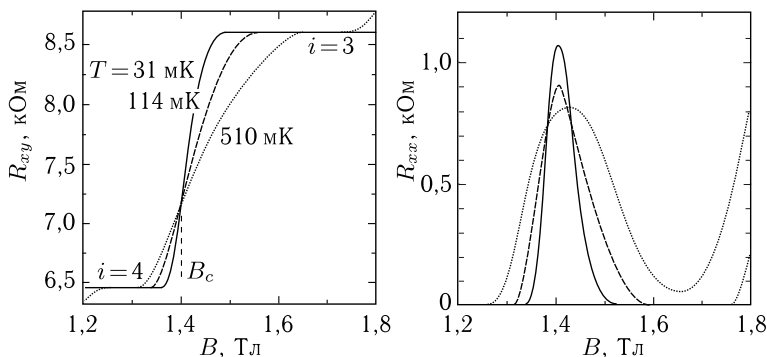


Рис. 10.7. Поперечное R_{xy} и продольное R_{xx} сопротивления при разных температурах на гетероструктуре $\text{Al}_x\text{Ga}_{1-x}\text{As}-\text{Al}_{0,33}\text{Ga}_{0,67}\text{As}$ со значением $x = 0,85\%$. Критическое магнитное поле перехода 4–3, определенное по точке пересечения кривых $R_{xy}(T)$, равно $B_c = 1,40$ Тл [8]

Согласно формуле (10.43), когда управляющий параметр принимает критическое значение, скейлинговая переменная

$$u = \delta x / T^{1/z\nu} \quad (10.44)$$

тождественно равна нулю при всех температурах и

$$R_{uv}(x_c, T) = \text{const}. \quad (10.45)$$

Следовательно, *сепаратрисса* (10.45) *должна быть горизонтальна, а все изотермы* $R_{uv}(x, T = \text{const})$ *должны пересекаться в одной точке* $x = x_c$. Это первое условие применимости теории, и оно выполнено для кривых $R_{uv}(T)$ на рис. 10.7.

Сосредоточимся сначала на кривых $R_{xy}(T)$. Вблизи точки пересечения все кривые $R_{xy}(T)$ можно, разложив в ряд, заменить на прямые $(\partial R_{xy}/\partial B)_{B_c}(B - B_c)$. При изменении наклона прямых с $(\partial R_{xy}/\partial B)_{B_c}$ на $(\partial R_{xy}/\partial B)_{B_c}/T^\kappa$, где $\kappa = 1/z\nu$, все прямые должны слиться в одну. Возможность подобрать такое значение κ , чтобы выполнялось соотношение

$$(\partial R_{xy}/\partial B)_{B_c}(T)/T^\kappa = \text{const}, \quad (10.46)$$

это второе условие применимости теории. Выбор κ осуществляется путем построения графика $(\partial R_{xy}/\partial B)$ от T в двойном логарифмическом масштабе (см. рис. 10.8).

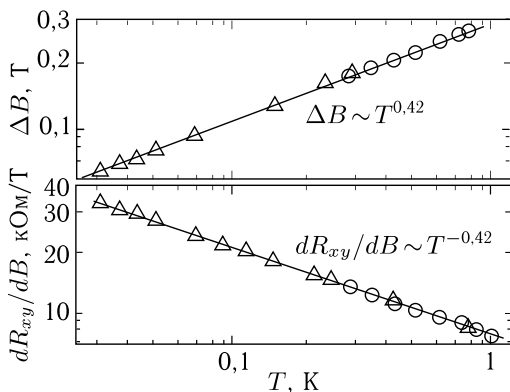


Рис. 10.8. Определение критического индекса $\kappa = 1/z\nu$ для перехода 4–3 по совмещению пересекающихся кривых $R_{xy}(T)$ (внизу) и по ширине пика $R_{xx}(T)$ (вверху). Данные получены на том же образце, что и на предыдущем рисунке, в криостате растворения (треугольники) и в криостате с жидким ^3He (кружки) [8]

Обратимся теперь к продольному сопротивлению R_{xx} , к которому формально уравнение (10.43) применимо в той же мере, что и к R_{xy} . В предыдущей главе при обсуждении экспериментальных кривых, например, на рис. 9.7 и 9.21, неявно подразумевалось, что пики R_{xx} указывают точно на точки переходов. Теперь видно, что это не так, потому что высота пиков зависит от температуры, т.е. точка максимума не удовлетворяет условию (10.45). Причина смещения максимума относительно точки перехода в том, что квантовый эффект Холла принципиально может наблюдаться только на образцах конечных размеров, а практически эти размеры обычно довольно малы. Поэтому мы опять сталкиваемся со скейлингом в системе ограниченных размеров («finite-size scaling»), причем на этот раз это не только размер вдоль оси мнимого времени, но и размеры в обычном пространстве.

Пики в продольном сопротивлении R_{xx} указывают на тенденцию к расходимости этой величины на переходе. Известно, что физические величины, расходящиеся на термодинамическом переходе в бесконечной системе, в системах с конечными размерами имеют максимум при близких, но, вообще говоря, других значениях управляющего параметра. Поэтому физические величины, расходящиеся на переходе, более чувствительны к конечности размеров реальных систем. Это делает процедуру, использованную при анализе функций $R_{xy}(B)$, ненадежной и неубедительной применительно к функциям $R_{xx}(B)$.

Тем не менее, данные по продольному сопротивлению можно использовать для скейлингового анализа, если воспользоваться интегральным свойством функций $R_{xx}(B)$ в окрестности перехода — полушириной ΔB пика, определенной по некоторому алгоритму. На рис. 10.8 величина ΔB определена как расстояние между двумя максимумами производной $(\partial R_{xx}/\partial B)$.

Как видно из рис. 10.8, анализ обоих семейств функций дал одинаковые значения критического индекса $\kappa = 0,42$. Это значение получалось неоднократно на гетероструктурах из различных материалов. Однако, вопреки ожиданиям, оно оказалось не универсальным: в целом ряде экспериментов на разных гетероструктурах наблюдались и другие значения κ , в интервале от 0,2 до 0,8. Поскольку скейлинговые соотношения и критические индексы обычно универсальны, наблюдаемый разброс требует объяснений. Одна из возможных причин в том, что плотность состояний $g(\varepsilon)$ в окрестности энергии ε_c зависит от особенностей случайного потенциала: потенциал может быть длинно- или короткопериодным, статистически симметричным или не симметричным относительно среднего значения и т.п. Работа [8], графики из которой приведены в этом параграфе, была как раз посвящена выяснению влияния характера случайного потенциала на κ .

В заключение два замечания относительно экспериментальной процедуры определения κ .

Хотя с точки зрения теории горизонтальность сепаратрисы (10.45) является неременным условием, на практике кривая $R_{uv}(x = x_c, T)$ может оказаться не совсем горизонтальной, а изотермы могут пересекаться не совсем в одной точке, что затрудняет точное определение значения x_c . Однако, в реализованной на рис. 10.8 процедуре (10.46) определения κ значение x_c в явном виде не используется. Достаточно *знать, что переход существует*, и предполагать, что наблюдаемые в эксперименте отклонения вызваны вторичными причинами.

Формула (10.41) утверждает, что внутри критической области скейлинг реализуется для всей функции, а не только для ее линейной части. Однако, при этом первые производные должны заведомо совпасть. Поэтому самый удобный и точный способ произвести скейлинг состоит в том, чтобы сначала определить критический индекс κ по первой производной, а потом сравнить сами функции и оценить размеры критической области.

Список литературы

1. Гантмахер В. Ф., Долгополов В. Т. Квантовые фазовые переходы «локализованные–делокализованные электроны» // УФН **178**, 3 (2008).
2. Sachdev S. Quantum Phase Transitions. Cambridge University Press, Cambridge 2000.
3. Sondhi S. L., Girvin S. M., Carini J. P., and Shahar D. Continuous quantum phase transitions // Rev. Modern Phys. **69**, 315 (1997).
4. Vojta T. Quantum phase transitions // In: K. H. Hoffmann and M. Schreiber (Eds): Computational Statistical Physics. Springer, Berlin (2002); cond-mat/0010285.
5. Goldenfeld N. Lectures on Phase Transitions and the Renormalization Group. Addison-Wesley, Reading, 1992.
6. Гантмахер В. Ф., Долгополов В. Т. Квантовый фазовый переход сверхпроводник–изолятор // УФН **180**, 3 (2010).
7. Huckestein B. Scaling theory of the integer quantum Hall effect // Rev. Mod. Phys. **67**, 357 (1995).
8. Wanli Li, Csathy G. A., Tsui D. C., Pfeiffer L. N., and West K. W. // Phys. Rev. Lett. **94**, 206807 (2005).

Приложение А

ЭЛЕМЕНТЫ ТЕОРИИ ПЕРКОЛЯЦИИ ¹⁾

Как любая ветвь математики, теория перколяции находит применение в самых разнообразных областях человеческой деятельности. Мы в этой главе изложим справочным образом основные положения и результаты этой теории, имея в виду специфический круг проблем транспорта в неоднородной среде, затрагиваемых в этой книге. Представляя среду в виде дискретной решетки, сформулируем два простейших типа задач. Можно выборочно случайным образом красить (открывать) узлы решетки, считая долю крашенных узлов x основным независимым параметром и полагая два крашенных узла принадлежащими одному кластеру, если их можно соединить непрерывной цепочкой соседних крашенных узлов. Такие вопросы как среднее число узлов в кластере, распределение кластеров по размерам, появление бесконечного кластера и доля входящих в него крашенных узлов, составляют содержание задачи узлов. Можно также выборочно красить (открывать) связи между соседними узлами и считать, что одному кластеру принадлежат узлы, соединенные цепочками открытых связей. Тогда те же самые вопросы о среднем числе узлов в кластере и т. д. составляют содержание задачи связей.

Когда все узлы (или все связи) закрыты, решетка является моделью изолятора. Когда они все открыты, и по проводящим связям через открытые узлы может идти ток, то решетка моделирует металл. При каком-то критическом значении $x = x_c$ произойдет перколяционный переход, являющийся геометрическим аналогом перехода металл–изолятор. Теория перколяции важна именно в окрестности перехода. Вдали от перехода достаточно аппроксимации эффективной среды.

А.1. Аппроксимация эффективной среды ²⁾

Рассмотрим ортогональную изотропную решетку с периодом a размерности d , у которой все узлы соединены с ближайшими соседями

¹⁾ Элементарное математическое изложение теории перколяции можно найти в книгах [1, 2]; оно же, но применительно к проблемам физики твердого тела, имеется в книге [3]. Ряд физических приложений обсуждается также в обзоре [4].

²⁾ С этим вопросом можно познакомиться также в сборнике [5], в который включена обзорная статья С. Киркпатрика.

связями с проводимостями σ_{kl} , независимыми друг от друга. Для примера на рис. А.1 изображена квадратная решетка $d = 2$. Требуется по известной дисперсии значений σ_{kl} вычислить среднее значение σ_m , определяющее удельную проводимость сетки

$$\sigma = \sigma_m a^{2-d}. \quad (\text{А.1})$$

Предположим сначала, что все σ_{kl} одинаковые, $\sigma_{kl} = \sigma_m$. Подведем два электрода к двум соседним узлам решетки k и l и пропустим через них ток J . Чтобы вычислить ток через связь σ_{kl} , отнесем сначала

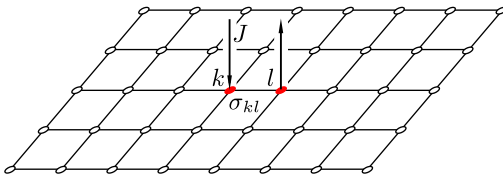


Рис. А.1. Для тока J , текущего через два соседних узла, k и l , квадратной решетки, проводимость kl -связи σ_{kl} зашунтирована проводимостью связей всей остальной решетки σ_{sh}

электрод l на бесконечность. Тогда ток J , втекающий через электрод k , растекается равномерно по z связям, подходящим к узлу k , так что через каждую связь идет ток J/z . (На квадратной решетке на рис. А.1 параметр $z = 4$, на трехмерной кубической решетке $z = 6$.) Теперь унесем на бесконечность электрод k , а ток J выведем через возвращенный на место электрод l . К узлу l этот ток стекается равномерно через все связи; в частности, через связь σ_{kl} идет такой же ток J/z в том же направлении, что и в предыдущем случае. Суперпозиция этих двух конфигураций дает интересующее нас распределение токов: токи J и $-J$ через электроды k и l и ток $J = 0$ на бесконечности. При этом ток через связь σ_{kl} равен

$$i_{kl} = \frac{2J}{z}, \quad (\text{А.2})$$

и из закона Ома следует, что

$$\frac{2J}{z\sigma_m} = \frac{J}{\sigma_m + \sigma_{sh}}, \quad (\text{А.3})$$

где в знаменателе справа стоит сумма проводимости самой kl -связи $\sigma_{kl} = \sigma_m$ и проводимости шунтирующей ее всей остальной сетки σ_{sh} , которую можно из соотношения (А.3) и определить:

$$\sigma_{sh} = \sigma_m \left(\frac{z}{2} - 1 \right). \quad (\text{А.4})$$

Соотношения (А.2)—(А.4) определяют распределение тока между непосредственной связью, соединяющей узлы k и l , и ее окружением. Удельная же проводимость определяется величиной σ_m через соотношение (А.1).

Теперь пусть все σ_{kl} различны. Тогда из закона Ома следует лишь, что для любой пары соседних узлов k, l

$$\frac{i_{kl}}{\sigma_{kl}} = \frac{J - i_{kl}}{\sigma_{sh}}, \quad i_{kl} = J \frac{\sigma_{kl}}{\sigma_{kl} + \sigma_{sh}}. \quad (\text{A.5})$$

Это выражение нужно усреднить по всем возможным парам. Схема усреднения называется аппроксимацией эффективной среды. В этой аппроксимации делаются два предположения: (i) в системе со случайными σ_{kl} по-прежнему справедливо соотношение (A.4) для величины σ_{sh} , поскольку шунтирование осуществляется большим числом связей, по которому реально всегда происходит усреднение (в этом смысл названия аппроксимации; иногда пользуются также названием «приближение среднего поля»); при этом подразумевается, что входящая в (A.4) величина σ_m по-прежнему определяет удельную проводимость в соответствии с соотношением (A.1); (ii) если подводить электроды последовательно к большому количеству пар узлов, то среднее значение тока $\langle i_{kl} \rangle$ определяется выведенным для однородного случая выражением (A.2)

$$\langle i_{kl} \rangle = \frac{2J}{z}. \quad (\text{A.6})$$

Воспользовавшись при усреднении выражения (A.5) этими двумя предположениями, получим окончательное выражение

$$\frac{2}{z} = \left\langle \frac{\sigma_{kl}}{\sigma_{kl} + \sigma_m(z/2 - 1)} \right\rangle. \quad (\text{A.7})$$

Для решения конкретной задачи остается подставить в соотношение (A.7) функцию распределения для величины σ_{kl} . Если $\sigma_{kl} = 1$ с вероятностью x и $\sigma_{kl} = 0$ с вероятностью $1 - x$, то ответ получается в виде прямой линии с наклоном, зависящим от числа ближайших соседей z :

$$\frac{2}{z} = x \frac{1}{1 + \sigma_m(z/2 - 1)}, \quad \sigma_m = \frac{x - 2/z}{1 - 2/z}. \quad (\text{A.8})$$

На рис. A.2 показаны получившиеся зависимости для квадратной ($d = 2$, $z = 4$) и кубической ($d = 3$, $z = 6$) решеток. Обращение функции $\sigma_m(x)$ в нуль означает перколяционный переход. Но именно в окрестности этого перехода аппроксимация эффективной среды не работает, потому что не оправданы оба предположения, лежащие в ее основе. Поэтому настоящее критическое значение x_c , которое называется перколяционным порогом, находится не в точке пересечения $\sigma_m(x)$ с осью абсцисс, а левее. Это станет видно из данных, собранных в следующем параграфе.

Заметьте. На том же рис. A.2 приведены графики функции $\sigma_m(x)$ и для случая, когда σ_{kl} с вероятностью $1 - x$ меняет свое значение с 1 не на 0, а на $1/2$. Эти графики уже не прямые линии. При таких значениях σ_{kl} перколяционного перехода нет и аппроксимация эффективной среды применима во всем интервале изменения x .

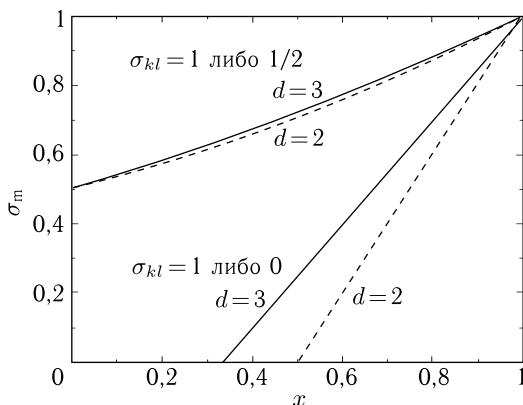


Рис. А.2. Проводимость квадратной и кубической решеток в аппроксимации эффективной среды, когда доля $(1-x)$ связей разорвана (две нижние кривые), или когда они имеют проводимость, в два раза меньшую проводимости остальных связей (две верхние кривые)

А.2. Перколяционные пороги

В окрестности перколяционного перехода, где аппроксимация эффективной среды не работает, получить точные результаты аналитическими методами удается довольно редко. В общем виде решения известны для одномерной решетки $d=1$ и бесконечномерной решетки Бете $d=\infty$, при $d=2$ есть решения для некоторых простых решеток. Чаще всего задачи решаются численно методами компьютерного моделирования. Это относится, в частности, к определению перколяционных порогов x_c , которые зависят от размерности d и от симметрии решетки. Для наиболее распространенных 2D- и 3D-решеток значения $x_c^{(s)}$ для задачи узлов и $x_c^{(b)}$ для задачи связей приведены в табл. А.1.

Приведенные значения x_c удовлетворяют некоторым эмпирическим закономерностям. В задаче связей произведение I_b числа связей z , соединяющих каждый узел со всеми ближайшими соседями, и критического значения $x_c^{(b)}$ зависит только от размерности и приблизительно равно $I_b = zx_c^{(b)} = d/(d-1)$. В задаче узлов в соответствующий инвариант $I_s = fx_c^{(s)}$ вместо z входит коэффициент f . Он равен доле объема, занятого шарами одинакового радиуса, расположенными во всех узлах решетки и касающимися друг друга (для 2D-решеток вместо шаров подразумеваются круги, а вместо объема — площадь). Как видно из таблицы, из 2D-решеток этот коэффициент максимален для треугольной решетки, а из 3D-решеток — для гранецентрированной. Именно поэтому эти решетки называются плотно упакованными.

Значения эмпирических инвариантов I_s и I_b показывают, что фундаментальным параметром в перколяционных задачах является не

Таблица А.1

	Задача узлов			Тип решетки	Задача связей		
	$x_c^{(s)}$	f	$I_s = f x_c^{(s)}$		$x_c^{(b)}$	z	$I_b = z x_c^{(b)}$
$d = 2$	0,59	0,79	0,47	Квадратная	0,5*	4	2,0
	0,5*	0,91	0,46	Треугольная	0,35*	6	2,1
	0,7	0,61	0,43	Медовые соты	0,65*	3	2,0
$d = 3$	0,31	0,52	0,16	Простая кубическая	0,25	6	1,50
	0,25	0,68	0,17	Объемноцентрированная	0,18	8	1,44
	0,20	0,74	0,15	Гранецентрированная	0,12	12	1,44
	0,43	0,34	0,15	Структура алмаза	0,39	4	1,56
Звездочкой * отмечены значения x_c , вычисленные аналитически.							

симметрия решетки и не число ближайших соседей, а размерность пространства. Пусть в крашенных (открытых) узлах (рис. А.3) находятся проводящие шары диаметром, обеспечивающим касание шаров на соседних узлах. Перколяция происходит при появлении уходящих на бесконечность цепочек касающихся шаров. Из значений инвариантов следует, что перколяционный переход на плоскости ($d = 2$) происходит, когда около 45 % ее становятся проводящими, а при переходе в пространстве ($d = 3$) проводящими являются примерно его 16 %.

Задачу узлов можно несколько изменить, признав в качестве «ближайших соседей», помимо истинно ближайших, и узлы второго, третьего и т. д. слоев, окружающих исходный узел. Число соседей, с которым связан данный узел, растёт, условия присоединения его к кластеру облегчаются и, как следствие, уменьшается критическое значение x . На рис. А.3 изображена квадратная решетка. Открытые узлы отмечены крестиками. Если считать ближайшими соседями три слоя узлов (рис. А.3, а), то из 10 изображенных открытых узлов в один кластер попадут 4, а если расширить область «сильного взаимодействия» до четырех слоев (рис. А.3, б), то в одном кластере окажутся уже 7 узлов из 10. Данные для трехмерных решеток приведены в табл. А.2. В ней собраны значения критических концентраций для трех разных решеток при включении узлов из одного, двух или трех слоев в число ближайших.

В последней строке таблицы выписаны критические значения величины zx :

$$zx \approx \frac{4\pi}{3} \left(\frac{r}{a}\right)^3 x = \frac{4\pi}{3} r^3 N. \tag{А.9}$$

Здесь r — радиус сферы, охватывающей все слои связанных с центром узлов, a^3 — объем элементарной ячейки решетки, т. е. объем, приходящийся на один узел, а $N = x/a^3$ — количество окрашенных узлов на единицу объема, т. е. концентрация, не связанная с параметрами конкретной решетки. Приблизительное равенство (А.9) выполняется

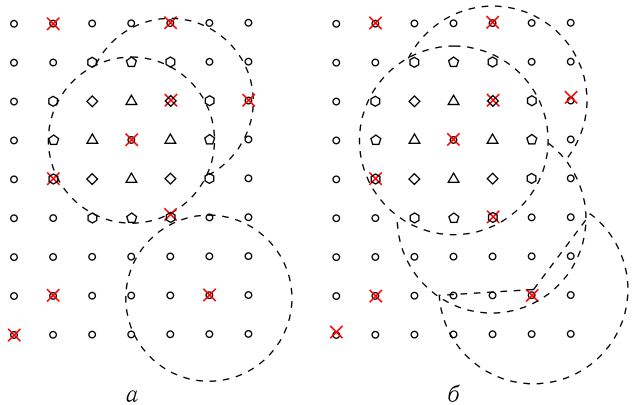


Рис. А.3. Квадратная решетка, в которой открытые узлы отмечены крестиками, а окружности охватывают узлы, связанные с узлом, находящимся в центре. Слева: каждый открытый узел входит в один кластер с открытыми узлами из трех слоев ближайших соседей (трех-, четырех- и пятиугольники). Справа: из четырех слоев (те же и шестиугольники)

тем точнее, чем больше число слоев ν и радиус r . Поэтому при достаточно больших r величина $zx_c^{(s\nu)}$ в последней строке таблицы уже не должна зависеть от конкретной решетки. Компьютерное моделирование подтверждает это. Оно свидетельствует также, что при $\nu \rightarrow \infty$ она стремится к пределу, равному $B_c^{(3)} = 2,7$.

Таблица А.2

	Простая кубическая			Объемноцентри рованная			Гранецентриро ванная		
Число слоев ν	1	2	3	1	2	3	1	2	3
Число соседей z	6	18	26	8	14	26	12	18	42
$x_c^{(s\nu)}$	0,31	0,14	0,10	0,25	0,175	0,095	0,195	0,14	0,06
$zx_c^{(s\nu)}$	1,84	2,45	2,52	1,94	2,45	2,47	1,84	2,45	2,52

Универсальность предела величины $zx_c^{(s\nu)}$ относительно разных решеток не случайна. При большом r можно без существенных последствий сместить окрашенные узлы из точно зафиксированных точек узлов решетки на расстояния $s \ll a \ll r$. Тем самым мы перейдем к задаче о перколяции в системе случайных узлов, которая эквивалентна задаче узлов на решетке при достаточно большом ν . Для перколяции на случайных узлах

$$\frac{4\pi}{3} r^3 N_c = B_c^{(3)} = 2,7; \qquad \pi r^2 N_c = B_c^{(2)} = 4,4.$$

(А.10)

Выражение в левой части второго равенства написано для 2D-решеток по аналогии с соотношением (A.9), а численное значение получено в результате компьютерного моделирования.

В соотношениях (A.10) предполагалось, что задано максимальное расстояние r , на котором узлы оказываются связанными (радиус взаимодействия). Тогда эти соотношения определяют критическую концентрацию N_c . Но возможна и иная постановка вопроса. Если задать концентрацию N , то соотношения (A.10) определяют перколяционный радиус r_c , т. е. минимальный радиус взаимодействия, обеспечивающий перколяцию.

Особый класс перколяционных задач составляют *континуальные задачи*. Зададим в пространстве размерности d случайную непрерывную функцию $U_{\min} \leq U(\mathbf{r}) \leq U_{\max}$ со средним значением $\overline{U(\mathbf{r})} = 0$ и со статистическими свойствами, инвариантными относительно преобразования $U \rightarrow -U$. Назовем крашеным кластером всякую связную область, где удовлетворено неравенство $U(\mathbf{r}) < u$, а белым ту, где справедливо обратное неравенство $U(\mathbf{r}) > u$. Так определенные кластеры аналогичны тем, которые фигурируют в решеточных задачах. Обозначим через $S_1(u)$ суммарный объем всех крашенных кластеров, нормированный на единичный объем, а через $S_2(u)$ — так же нормированный объем всех белых кластеров, $S_1 + S_2 = 1$. Будем смещать уровень u от $U_{\min}$ вверх. При достаточно малых $u \gtrsim U_{\min}$ крашенных кластеров мало, и они имеют малые размеры, но существует один белый бесконечный кластер. С ростом u средние размеры крашенных кластеров растут. Появление бесконечного крашеного кластера при $u = u_{c1}$ и $S_1(u) = S_{c1}$ означает перколяционный переход. На противоположном конце интервала возможных значений уровня u , при $u \lesssim U_{\max}$, существует крашенный бесконечный кластер. С уменьшением u в этой области растут число и средние размеры белых кластеров, пока при некоторых $u = u_{c2}$ и $S_2(u) = S_{c2}$ не появится бесконечный белый кластер. Из статистической симметрии случайного потенциала $U(\mathbf{r})$ следует, что $S_{c1} = S_{c2}$. Конкретные значения S_{c1} и S_{c2} зависят от размерности d . Для нас наиболее интересны случаи $d = 3$ и $d = 2$, на которых мы кратко остановимся.

В трехмерном случае, $d = 3$, компьютерное моделирование привело к значениям $S_{c1} = S_{c2} = 0,17$. В интервале значений уровня $u_{c1} < u < u_{c2}$ имеются одновременно и крашенный и белый бесконечные кластеры. Такая задача появляется при изучении локализации классического электрона: если $U(\mathbf{r})$ — потенциальная энергия электрона, то крашенные кластеры представляют собой области, классически доступные для электрона с энергией u . Перколяционный переход при $u = u_{c1}$ означает, что только электрон с энергией $u > u_{c1}$, двигаясь по законам классической механики, может уйти на бесконечность.

Переходя к двумерному пространству $d = 2$, воспользуемся для наглядности «географической терминологией», полагая, что $U(\mathbf{r})$ — это высота на местности точки с координатой $\mathbf{r}$. Локальные максимумы U_+

и минимумы U_- этой функции — это высоты холмов и глубины впадин. Представим себе, что на местности можно менять уровень воды. Пока уровень низок, крашенные кластеры — это озера вблизи точек U_- , а береговые линии озер — это эквипотенциальные линии потенциала $U(\mathbf{r})$, охватывающие точки U_- . При низком уровне воды пересечь всю территорию можно по суше. При постепенном подъеме воды ее уровень будет достигать локальных седловых точек функции $U(\mathbf{r})$, и будет происходить слияние озер. В обратном предельном случае, при очень высоком уровне воды, из нее выступают лишь островки вокруг вершин U_+ (белые кластеры), а пересечь всю территорию можно только по воде. Теперь эквипотенциальные линии — это периметр островов. Поэтому существует некий промежуточный уровень воды, при котором береговая линия простирается через всю территорию, но в некоторых местах уровень воды оказывается точно на высоте перевала в седловой точке, так что ширины водной и сухой перемычек стягиваются в точку.

Бесконечный кластер статистически изотропен. Если он существует, то территорию можно пересечь по нему в любом направлении. С другой стороны, если при каком-то уровне воды территорию можно пересечь по суше, то при этом ее заведомо нельзя пересечь в перпендикулярном направлении по воде. Это означает, что при $d = 2$ одновременно не могут существовать два бесконечных кластера. Если добавить статистическую симметрию относительно преобразования $U \rightarrow -U$, то получим, что

$$S_{c1} = S_{c2} = 0,5. \quad (\text{A.11})$$

При отсутствии такой симметрии перколяционный порог может сместиться, так что равенство (A.11) нарушится. Но запрет на одновременное существование двух бесконечных кластеров сохранится, так что вместо (A.11) будем иметь лишь $S_{c1} + S_{c2} = 1$.

Интересно сравнить получившиеся значения S_c для двух размерностей с имеющим тот же физический смысл инвариантом I_s в задаче узлов на регулярных решетках — см. таблицу А.1. В обоих случаях значения S_c несколько больше, и если при $d = 3$ разница лежит на границе погрешности, то при $d = 2$ она не вызывает сомнений. Решетка вносит корреляции и элементы регулярности в структуру случайной функции. Это сдвигает точку появления бесконечного кластера.

А.3. Окрестность перколяционного перехода

Для исследования окрестности перколяционного перехода введем несколько важных функций концентрации открытых узлов x (для определенности мы будем говорить здесь о задаче узлов). Пусть n_s — это число кластеров из s узлов, приходящихся на один узел решетки,

так что при $x < x_c$, когда бесконечный кластер отсутствует,

$$\sum_s sn_s = x, \quad x < x_c. \quad (\text{A.12})$$

Если же $x > x_c$, то некоторая часть $P(x)$ открытых узлов входит в бесконечный кластер:

$$P(x) = \begin{cases} x - \sum_s sn_s, & x \geq x_c, \\ 0, & x < x_c. \end{cases} \quad (\text{A.13})$$

Функция $P(x)$ называется мощностью бесконечного кластера.

Определим корреляционную функцию $q(r)$ как вероятность того, что узел на расстоянии r от открытого узла, принадлежащего конечному кластеру, тоже открыт и принадлежит тому же кластеру. Очевидные свойства этой функции таковы: $q(0) = 1$, $q(a) = x$ (a — период решетки); поскольку по определению кластер конечный, то $q(r) \rightarrow 0$ при $r \rightarrow \infty$. Среднее число узлов $S(x)$ в конечном кластере, к которому принадлежит исходный открытый узел, равно $\sum_r q(r)$; здесь суммирование подразумевается по всем узлам решетки.

Величину $S(x)$ можно выразить и непосредственно через распределение кластеров по размерам. Вероятность того, что произвольный узел принадлежит кластеру с числом узлов s , равна sn_s , а вероятность того, что он принадлежит к какому-то из конечных кластеров, есть $\sum_s sn_s$. Поэтому отношение $w_s = sn_s / \sum_s sn_s$ есть вероятность того, что кластер, к которому принадлежит произвольно выбранный открытый узел, содержит s узлов. Сделав такой выбор большое количество раз, получим среднюю величину

$$S(x) = \sum_r q(r) = \sum_s sw_s = \frac{\sum_s s^2 n_s}{\sum_s sn_s}. \quad (\text{A.14})$$

Функция $S(x)$ называется средним размером кластера. (Заметьте: подразумевается не линейный размер, а среднее число узлов конечного кластера.)

Линейный размер, связанный с распределением открытых узлов по конечным кластерам, определим соотношением

$$\xi^2 = \frac{\sum_r r^2 q(r)}{\sum_r q(r)} = \frac{\sum_r r^2 q(r)}{S(x)}. \quad (\text{A.15})$$

Функция $\xi(x)$ называется корреляционной длиной. По существу, это средний линейный размер среднего кластера, «типичного» для данного x . Важно, что функция $\xi(x)$ определена по обе стороны от перколяционного порога. При $x > x_c$ конечные кластеры располагаются в «дырах» бесконечного кластера. Поэтому в этой области корреляционную

длину обычно интерпретируют как средний размер дыр бесконечного кластера. По обе стороны от порога определена также и функция $S(x)$.

Функция $P(x)$ обращается на перколяционном пороге в нуль, функции $S(x)$ и $\xi(x)$ на пороге обращаются в бесконечность. Основной постулат, лежащий в основе теоретического описания перколяционного перехода, заключается в том, что при приближении к порогу все эти функции изменяются как степени расстояния до порога:

$$\left. \begin{aligned} P(x) &\propto (x - x_c)^\beta, & x > x_c, \\ S(x) &\propto |x - x_c|^{-\gamma} \\ \xi(x) &\propto |x - x_c|^{-\nu} \end{aligned} \right\} \begin{array}{l} \text{Показатели } \gamma \text{ и } \nu \text{ одинаковые} \\ \text{по обе стороны порога.} \end{array} \quad (\text{A.16})$$

Эти функциональные зависимости неоднократно проверялись и продолжают проверяться экспериментально. Компьютерное моделирование не только подтвердило наличие степенной функциональной зависимости, но и продемонстрировало универсальность показателей степени β , γ и ν : они зависят только от размерности пространства d , но не зависят ни от симметрии решетки, ни от типа задачи. Это характерное свойство теории фазовых переходов. Там аналогичные показатели степени называются критическими индексами; они тоже универсальны и тоже зависят только от размерности. Поэтому на сегодняшний день считается твердо установленным, что с точки зрения адекватного математического описания перколяционный переход аналогичен фазовому переходу второго рода. При этом:

- доля открытых узлов (связей) x играет роль температуры;
- мощность бесконечного кластера $P(x)$ аналогична параметру порядка;
- функция $\xi(x)$ есть локализационная длина и тут и там;
- средний размер кластера $S(x)$ следует сравнивать с термодинамической функцией, например, с восприимчивостью при магнитном переходе.

Значения критических индексов вблизи перколяционного перехода в двумерных системах, $d = 2$, получены аналитически, а для размерности $d = 3$ — численно. Они приведены в табл. А.3.

Таблица А.3

		$d = 2$	$d = 3$
Мощность бесконечного кластера P	β	5/36	0,417
Локализационная длина ξ	ν	4/3	0,875
Среднее число узлов в кластере S	γ	43/18	1,795

А.4. Пример: электропроводность сильно неоднородной среды

В заключение этого приложения разберем конкретный пример использования теории перколяции для решения физической задачи. Сам пример этот тоже модельный, но он демонстрирует существо подхода. Итак, рассмотрим в трехмерном пространстве, $d = 3$, простую кубическую решетку с периодом a , связи в которой имеют сопротивления в экспоненциально большом интервале значений

$$R = R_0 \exp u, \quad 0 \leq u \leq u_0, \quad u_0 \gg 1. \quad (\text{A.17})$$

Пусть u — это случайная величина, с вероятностью $F(u)$ принимающая любые значения из разрешенного интервала. Выберем некоторое, достаточно малое значение u' из этого интервала и все связи с сопротивлениями от R_0 до $R_0 \exp u'$ сохраним, а все связи с большими сопротивлениями временно разорвем. Доля открытых (сохраненных) связей равна

$$x = \int_0^{u'} F(u) du, \quad \int_0^{u_0} F(u) du = 1. \quad (\text{A.18})$$

Пусть $x < x_c$. Это означает, что набор включенных сопротивлений не может обеспечить конечную проводимость решетки. В этот набор нужно добавить какое-то количество связей с большими сопротивлениями так, чтобы достичь порога $x = x_c$. (Напомним, что, согласно табл. А.1, в задаче связей для простой кубической решетки $x_c = 0,25$). Конечно, добавлять будем самые малые сопротивления из оставшихся, постепенно увеличивая u' в уравнении (А.18). При достижении порога соответствующее значение u' обозначим через u_c . Пусть для определенности все значения u из разрешенного интервала равновероятны, так что функция распределения $F(u) = \text{const}$:

$$F(u) = \begin{cases} 0, & u > u_0; \\ 1/u_0, & u \leq u_0. \end{cases} \quad (\text{A.19})$$

Тогда критическое значение параметра u равно

$$u_c = x_c u_0. \quad (\text{A.20})$$

Включенные на последнем этапе сопротивления $R_0 \exp u_c$ соединили большие конечные кластеры в один бесконечный. Они включены последовательно со всей совокупностью остальных сопротивлений и при этом они больше их всех по величине. Поэтому удельное сопротивление решетки контролируется именно этими сопротивлениями, и из-за них оно пропорционально $\exp u_c$.

Поскольку ток должен идти по связям, принадлежащим бесконечному кластеру, следует немного превысить порог до значения

$u'' = u_c + \Delta u$, чтобы мощность бесконечного кластера $P(x)$ стала отлична от нуля. Тогда из соотношений (А.19) и (А.20) следует, что $x - x_c = \Delta u/u_0$ и что корреляционная длина $\xi = a(u_0/\Delta u)^\nu$.

Для того чтобы оценить величину Δu , в частности убедиться, что $\Delta u \ll u_0$, обсудим распределение тока, протекающего по бесконечному кластеру. Большую часть бесконечного кластера вблизи порога составляют «бывшие» конечные кластеры, подсоединенные к основной части через одну открытую связь. Это хорошо видно на рис. А.4,

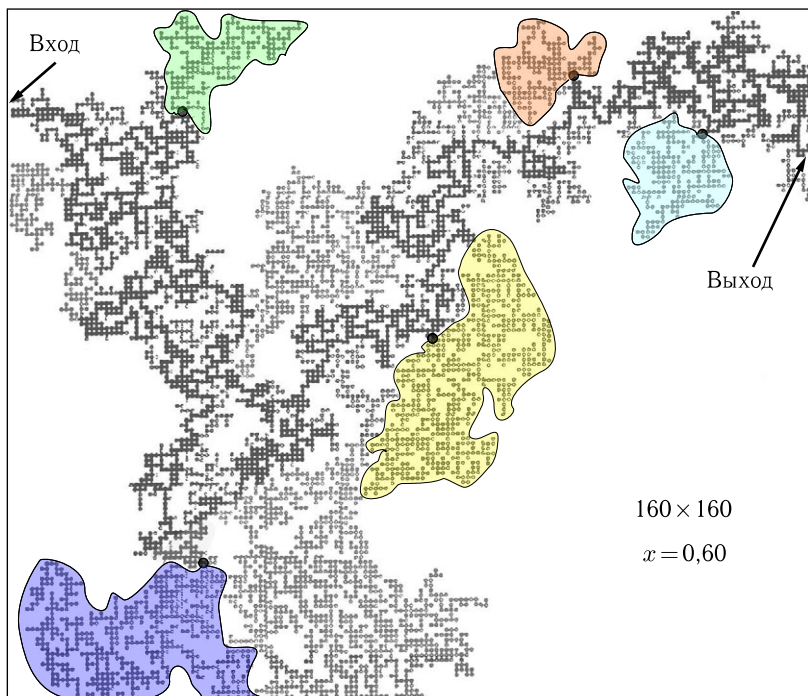


Рис. А.4. Экспериментальная реализация бесконечного кластера вблизи порога перколяции, полученная компьютерным моделированием задачи узлов на решетке 160×160 [2]. Показаны только открытые узлы, принадлежащие бесконечному кластеру. Из них темные точки — это узлы, принадлежащие токонесущему остоу. Любая из светлых точек соединена с токонесущим остовом только через один открытый узел — светлые точки входят в «мертвые концы». Самые крупные мертвые концы выделены серым фоном, а места их присоединения к токонесущему остоу отмечены черными кружками

где изображена экспериментальная реализация бесконечного кластера вблизи порога перколяции, полученная компьютерным моделированием задачи узлов на решетке 160×160 при значении $x = 0,6$. Кластер считается бесконечным, поскольку он соединяет противоположные (левую

и правую) грани решетки. Части кластера, выделенные серым фоном, подсоединены к основной части через один открытый узел каждая. Места их подсоединения условно отмечены черными кружками. Ток через них идти не может: чтобы шел ток, нужны по крайней мере две точки подсоединения участка к основной цепи. Поэтому эти участки называются мертвыми концами. Кроме выделенных на рисунке крупных мертвых концов есть еще много мелких. Часть бесконечного кластера, оставшаяся после исключения мертвых концов, является токонесящей. Входящие в нее узлы нарисованы на рис. А.4 черными точками.

В белых частях квадрата на рис. А.4 имеется та же средняя плотность открытых узлов, что и в областях, покрытых бесконечным кластером. Они не показаны, потому что все принадлежат различным конечным кластерам. Явная асимметрия представленной реализации бесконечного кластера объясняется тем, что из-за близости к порогу корреляционная длина ξ больше стороны квадрата: $\xi > 160$. На очень большой решетке токонесящую часть бесконечного кластера можно себе представить в виде сетки токовых каналов с масштабом, равным длине корреляции ξ (рис. А.5). В этой токонесящей сетке от граничного значения u'' зависят и размер ее ячейки ξ , и сопротивление между двумя ее узлами $R_\xi \approx e^{u_c + \Delta u}$. Удельное сопротивление решетки, представляемое как сопротивление токонесящей сетки, равно (см., в частности, (А.1))

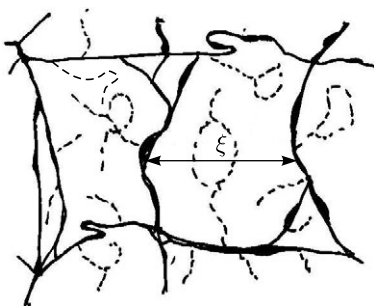


Рис. А.5. Структура токонесящей части перколяционного кластера

$$\rho = R_\xi \xi = R_0 e^{(u_c + \Delta u)a} \left(\frac{u_0}{\Delta u} \right)^\nu = R_0 a u_0^\nu e^{u_c} \frac{e^{\Delta u}}{(\Delta u)^\nu}. \quad (\text{А.21})$$

Дробь в правой части этого выражения имеет минимум при

$$\Delta u = \nu = 0,875 \approx 1 \ll u_0. \quad (\text{А.22})$$

Это означает, что подключение к сетке сопротивлений со значениями u в очень узком интервале $u_c < u < u_c + 1$ уменьшает сопротивление решетки за счет увеличения мощности бесконечного кластера P и уменьшения длины корреляции ξ . Затем дальнейшее включение сопротивлений становится неэффективным из-за того, что они уже много больше, чем $R_0 \exp u_c$, и не могут зашунтировать сложившуюся критическую токонесящую сетку. С точностью до численного множителя сопротивление рассматриваемой решетки равно

$$\rho = R_0 a u_0^\nu e^{u_c}. \quad (\text{А.23})$$

А.5. Контрольные вопросы и задачи

1. Узлы считаются связанными, если построенные вокруг них сферы радиуса r имеют хотя бы одну общую точку (перекрывающиеся сферы на случайных узлах). Найти долю объема, попадающего внутрь перекрывающихся сфер на пороге перколяции, т. е. при r , равном половине перколяционного радиуса, $r = r_c/2$.

$$1 - \exp(-B_c/8) = 0,29.$$

2. Пусть в одномерной цепочке x — вероятность заполнения узла, n_s — количество кластеров из s заполненных узлов, нормированное на полное число узлов:

- определить, чему равен перколяционный порог x_c ;
- написать выражение для n_s для любого s ;
- проверить формулу $\sum_s n_s s = x$;
- вычислить вероятность w_s того, что произвольный узел входит в s -кластер и средний размер кластера $S = \sum w_s s$;
- вычислить вероятность $g(r)$ того, что узел на расстоянии r шагов от занятого узла, тоже занят и принадлежит тому же кластеру (корреляционная функция).

См.: Stauffer D. Introduction to percolation theory, Taylor & Francis, 1985).

Список литературы

1. Stauffer D. Introduction to percolation theory. — Taylor & Francis, 1985.
2. Федер Е. Фракталы. Мир, 1991 [Перевод книги Feder J., Fractals, Plenum Press, 1988].
3. Шкловский Б. И., Эфрос А. Л. Электронные свойства легированных полупроводников, 1979, гл. 5.
4. Isichenko M. B. // Rev. Mod. Phys. **45**, 574 (1973).
5. Киркпатрик С. в сб. Теория и свойства неупорядоченных материалов (ред. В. Л. Бонч-Бруевич) С. 249. — Мир, 1977 [Перевод статьи Kirkpatrick S., Rev. Mod. Phys. **45**, 574 (1973)].

Приложение Б

ТУННЕЛЬНЫЕ ХАРАКТЕРИСТИКИ ¹⁾

Пусть два проводящих материала находятся в контакте друг с другом через слой изолятора, столь тонкий, что через него возможно туннелирование электронов. Эти два материала мы будем называть электродами, а изолятор между ними — контактом или туннельным промежутком. В равновесии ферми-уровни по обе стороны контакта одинаковы. Электрон с энергией ε относительно ферми-уровня слева от барьера после туннелирования, оказавшись справа, сохранит то же значение энергии. Равновесие является динамическим в том смысле, что потоки туннелирующих в обе стороны электронов отличны от нуля, но эти потоки равны между собой, так что полный ток через контакт равен нулю.

Обратите внимание на сходство и различие между туннелированием, о котором пойдет речь в этом Приложении, и прыжковой проводимостью (гл. 4). В обоих случаях речь идет о подбарьерном прохождении электронов через классически недоступную область. Но прыжковая проводимость осуществляется за счет квантовых переходов между локализованными состояниями. Туннельный же ток через контакт возникает благодаря квантовым переходам между делокализованными состояниями, находящимися по разные стороны от барьера.

В большинстве случаев сопротивление электродов много меньше сопротивления контакта. Поэтому приложенная к электродам разность потенциалов V полностью падает на туннельном промежутке. Равновесие нарушается, и через контакт течет туннельный ток J (рис. Б.1),

$$J(V) = A \int_{-\infty}^{\infty} g(\varepsilon - eV) g_1(\varepsilon) \left[f\left(\frac{\varepsilon - eV}{T}\right) - f\left(\frac{\varepsilon}{T}\right) \right] d\varepsilon, \quad (\text{Б.1})$$
$$A \propto S \exp[-2w(2m_{\text{ef}}U_{\text{ef}})^{1/2}/\hbar] = S e^{-2\kappa w}.$$

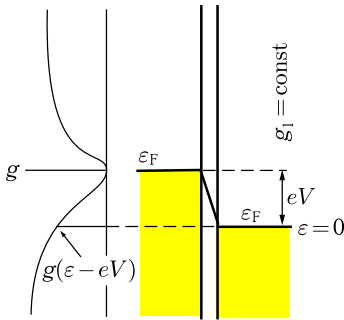
¹⁾ Существуют книги [1, 2], целиком посвященные туннельным явлениям; они изданы уже довольно давно. С последними достижениями в этой области можно познакомиться в обзоре [3]

Из всех множителей, входящих в коэффициент A , в уравнении (Б.1) выписаны только два: площадь S и содержащий параметры барьера экспоненциальный множитель, который является для A определяющим; здесь w — ширина туннельного барьера, U_{ef} — его эффективная высота; величина m_{ef} , имеющая размерность массы, определяется плотностями состояний по обе стороны барьера. Величина

$$\kappa = (2m_{\text{ef}}U_{\text{ef}})^{1/2}/\hbar$$

имеет размерность обратной длины.

Рис. Б.1. Энергетическая схема прохождения тока через туннельный контакт. Желтым цветом выделены состояния, заполненные при $T = 0$



Подынтегральное выражение описывает избыточный электронный поток электронов энергии ε слева направо, обусловленный тем, что ферми-распределение $f(x) = (\exp x + 1)^{-1}$ слева и справа от барьера сдвинуты на eV . Плотности состояний g и g_1 входят

суммножителями в подынтегральное выражение, потому что они определяют количество участвующих в процессе электронов и количество состояний, в которые возможно туннелирование.

Выражение (Б.1) сильно упрощается, если один из электродов — обычный металл с $g_1(\varepsilon) = \text{const}$, а температура столь низкая, что ферми-распределение можно считать ступенькой. Тогда

$$J(V) \propto g_1 \int_0^{eV} g(\varepsilon) d\varepsilon. \quad (\text{Б.2})$$

Если к постоянному напряжению добавить переменное $V_\omega \sin \omega t$, то появится переменный ток на частоте ω с амплитудой J_ω , пропорциональной производной dJ/dV :

$$J(V + V_\omega \sin \omega t) = J(V) + \frac{dJ}{dV} V_\omega \sin \omega t. \quad (\text{Б.3})$$

Этот распространенный экспериментальный прием называется модуляционной методикой. Из (Б.2) и (Б.3) следует

$$J_\omega \propto \frac{dJ}{dV} \propto g(eV). \quad (\text{Б.4})$$

Это и определяет возможность использования измерений туннельных токов для прямых измерений $g(\varepsilon)$.

Конечная температура $T \neq 0$ несколько усложняет формулу (Б.4):

$$J_{\omega} \propto dJ/dV \propto - \int_{-\infty}^{\infty} g(\varepsilon) d\varepsilon \frac{\partial}{\partial(eV)} f\left(\frac{\varepsilon - eV}{T}\right), \quad (\text{Б.5})$$

и процедуру извлечения функции $g(\varepsilon)$ из экспериментальных данных, но не меняет принципиальной возможности изучения особенностей плотности состояний вблизи ферми-уровня по вольт-амперным характеристикам туннельного контакта.

Заметьте. Реализация туннельного эксперимента гораздо сложнее, чем изложенная его принципиальная схема: нужно уметь надежно и воспроизводимо создавать между электродами туннельные промежутки толщиной 10–15 Å. Эта экспериментальная задача была решена только в 1960-х. Примененные к сверхпроводникам, туннельные эксперименты позволили непосредственно измерить сверхпроводящую щель, а затем привели к открытию эффекта Джозефсона и тем самым существенно повлияли на развитие всей физики низких температур.

Примененная к нормальным металлам, туннельная методика позволила непосредственно обнаружить кулоновскую щель в спектре локализованных состояний и минимум плотности состояний на ферми-уровне в спектре грязных металлов, а также проследить, как одно переходит в другое.

Пример выявления кулоновской щели в Si:B приведен на рис. Б.2. Была использована структура Pb–SiO₂–Si:B, в которой один электрод — сверхпроводник, а второй — классический полупроводник р-типа. В нулевом магнитном поле основным «нарушителем спокойствия» является сверхпроводящий свинец, а Si:B играет роль металлического контрэлектрода. Соответствующая кривая демонстрирует сверхпроводящую щель в Pb с максимумами плотности состояний по краям. Она, как и должно быть, симметрична относительно ферми-уровня $\varepsilon_F = 0$. Измерение сверхпроводящей щели выступает здесь уже как калибровочный эксперимент. Это как бы дополнительная демонстрация правильности методики. В поле 2 кЭ сверхпроводимость свинца полностью разрушена, и плотность состояний в нем больше не зависит от энергии. Теперь Pb становится контрэлектродом, а минимум на кривой dJ/dV вызван наличием параболической кулоновской щели в Si:B. Поле 2 кЭ слишком мало, чтобы хоть как-то повлиять на нее. Как видно из кривой, ширина щели примерно 1 мэВ ≈ 10 К (по 5 К в каждую сторону от ферми-уровня).

Кривые на рис. Б.2 получены при 1,15 К. Повторить их при более низких температурах затруднительно: из-за экспоненциального роста объемного сопротивления Si:B существенная часть прикладываемой к структуре разности потенциалов падает не на туннельном промежутке, а на электроде. Но несложная математическая обработка формулы

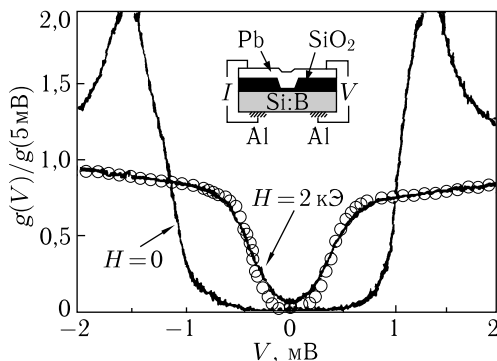


Рис. Б.2. Дифференциальная проводимость в зависимости от смещения на туннельном контакте Pb — SiO₂ — Si:B. Кружками показан результат пересчета кривой в поле $H = 2$ кЭ, убирающий влияние теплового уширения функции распределения [4]

(Б.5) позволяет обойтись без низкотемпературных измерений и извлечь из экспериментальной кривой $J_{\omega}(eV)$ (сплошные кривые на рис. Б.2) функцию $g(\varepsilon)$ (пустые кружочки). Видно, что пока температура мала, она лишь незначительно сглаживает экспериментальное отображение функции $g(\varepsilon)$.

Основная примесь в Si:B — акцепторы, и к нему полностью применима модель примесной зоны при слабом легировании, в рамках которой и была получена кулоновская щель (см. формулу (3.22) в гл. 3). Однако кулоновская щель наблюдается и в материалах совсем других типов, например, в ультратонких пленках или в аморфных пленках с химическим составом из двух элементов, один из которых — металл, а другой — неметалл.

На рис. Б.3 представлены туннельные характеристики структуры $\text{Ge}_{1-x}\text{Au}_x\text{-Al}_2\text{O}_3\text{-Al}$. Здесь Al является контрэлектродом, а $\text{Ge}_{1-x}\text{Au}_x$ — исследуемым материалом. Серия характеристик демонстрирует эволюцию спектра по мере уменьшения концентрации золота x . Минимум при больших x обусловлен взаимодействием электронов, диффундирующих в результате упругого рассеяния на статическом беспорядке. Это тот минимум на ферми-уровне в спектре грязных металлов, которого следует ожидать на основании теории, результаты которой изложены в последнем параграфе гл. 2. Но эти результаты получены в рамках теории возмущений, и предполагается, что поправки к функции $g(\varepsilon)$ малы. Поэтому кривые на рис. Б.3 не просто подтверждают наличие минимума в спектре, но и показывают эволюцию поправки по мере приближения к переходу металл-изолятор. При самой маленькой концентрации золота, при $x = 0,08$, функция $g(\varepsilon)$ обращается в нуль при $\varepsilon = 0$, а в окрестности этой точки является параболой. Это и есть кулоновская щель.

Заметьте. Модель примесной зоны в частично компенсированном полупроводнике, в рамках которой в гл. 3 была получена кулоновская щель, непосредственно к $\text{Ge}_{1-x}\text{Au}_x$ при малых x неприменима, потому что в этом материале нет четко определенных «доноров» и «акцепторов». Поэтому обнаружение кулоновской щели в $\text{Ge}_{1-x}\text{Au}_x$ указывает на то, что кулоновская щель может возникать в разных классах случайных потенциалов. Это замечание относится также и к экспериментам на аморфных сплавах $\text{Si}_{1-x}\text{Nb}_x$ и на ультратонких пленках Ве.

Таким образом, серия кривых на рис. Б.3 демонстрирует три важных результата: наличие минимума на ферми-уровне в спектре грязного металла, возникающего вследствие межэлектронного взаимодействия (гл. 2), наличие кулоновской щели в спектре сильно разупорядоченном изолятора (гл. 3) и переход от одной особенности в спектре к другой, сопровождающий переход металл–изолятор при изменении концентрации металла в сплаве (гл. 5).

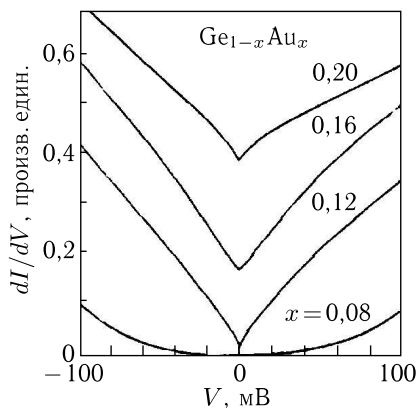


Рис. Б.3. Туннельные характеристики структуры $\text{Ge}_{1-x}\text{Au}_x\text{-Al}_2\text{O}_3\text{-Al}$ [5]

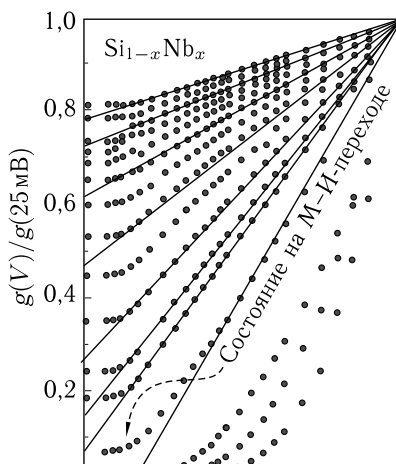


Рис. Б.4. Туннельные характеристики структуры $\text{Si}_{1-x}\text{Nb}_x\text{-Al}_2\text{O}_3\text{-Al}$ [6]

Такую же структуру и эволюцию спектра наблюдали на аморфных сплавах $\text{Si}_{1-x}\text{Nb}_x$, см. рис. Б.4, где тоже происходит переход металл–изолятор. В экспериментах на этих аморфных сплавах обнаружилось еще одно любопытное обстоятельство: несовпадение критического значения x_c , определенного в этих сплавах по измерениям статической проводимости, и значения $x_c^{(t)}$, определенного по обращению в нуль плотности состояний. Кривая с концентрацией x_c отмечена на рис. Б.4 надписью и стрелкой. Стрелка указывает, что при этой концентрации

на ферми-уровне остается конечная плотность состояний, а в нуль она обращается лишь на следующей кривой.

Из определения металла, данного в начале гл. 5, следует, что истинной критической концентрацией является x_c . На ферми-уровне изолятора вполне может быть конечная плотность состояний, если эти состояния были локализованы. Поэтому противоречия между измерениями проводимости и туннельными измерениями нет.

Вместе с тем возможна иная трактовка расхождения между x_c и $x_c^{(t)}$. Сделаем одно уточнение относительно измерений туннельного тока. В рассуждении, позволившем написать выражение для туннельного тока (Б.1), неявно предполагалось, что в обоих электродах электроны представляют собой системы невзаимодействующих квазичастиц. Предполагается, что полные электронные энергии и слева и справа равны $E = \sum \epsilon_i$, сумме энергий квазичастиц, и что в результате туннелирования полные энергии E с обеих сторон изменились точно на энергию протуннелировавшей квазичастицы. В системе взаимодействующих частиц полная энергия E зависит от их числа, $E = E(n)$. Тогда часть энергии электромагнитного поля тратится на изменение полных энергий электронных систем с изменившимся числом частиц. Во избежание недоразумений плотность состояний g_i , входящая в выражение (Б.1) и измеряемая в туннельном эксперименте, называется туннельной плотностью состояний.

Заметьте. Механизмы воздействия разных видов межэлектронного

взаимодействия на туннельную плотность состояний различны.

Вот, например, как воздействует на нее кулоновское взаимодействие туннелирующего электрона со всеми остальными. После туннелирования в обоих электродах возникает пространственная неоднородность зарядов, которая должна рассосаться. Чем больше время рассасывания, тем сильнее эта неоднородность препятствует самому туннелированию, проявляясь в виде уменьшения эффективной плотности состояний. Время рассасывания зависит, в частности, от характера движения электрона, от того, является это движение баллистическим или диффузионным. При диффузионном движении это время зависит от энергии электрона по сравнению с ферми-уровнем (см. гл. 2). Поэтому этот эффект может приводить не только к перенормировке туннельной плотности состояний, но и к изменению функциональной зависимости $g(\epsilon)$.

Аналогичное наблюдение было сделано на классическом полупроводнике с примесной зоной Si:B, где управляющим параметром является не состав сплава, а концентрация примесей n . Подробные измерения, представленные на рис. Б.5, показали, что и здесь $g(\epsilon_F)$ обращается в нуль уже в области изолятора, при $n \approx 0,9 n_c$ (критическая концентрация n_c определена по зависимости от n проводимости $\sigma(0)$). Расхождения в определении критической концентрации оказались порядка 10%.

Интересно, что на серии кривых с разными электронными концентрациями на рис. Б.5 минимум на кривой $g(V)$ при концентрации $n = n_c$, определенной как критическая по измерениям проводимости, оказался шире, чем по обе стороны от нее. Следовательно туннельные измерения тоже выделяют «транспортное» значение $n = n_c$, которое мы выше определили как истинное.

Специфическим объектом туннельных экспериментов являются ультратонкие пленки. Для них управляющим параметром может быть толщина пленок, а количественной характеристикой их транспортных свойств — величина сопротивления на квадрат при какой-то фиксированной температуре. Толщины пленок Ве, туннельные характеристики которых представлены на рис. Б.6, были в интервале 15–20 Å, указанные около кривых значения сопротивления на квадрат измерены при 50 мК. Туннельные структуры создавались дозированным окислением пленок на воздухе и последующим напылением на них слоя Ag. У самой тонкой пленки Ве в спектре видна кулоновская щель, причем плотность состояний в окрестности μ меняется линейно с энергией, как это и предсказывается формулой (3.23) в гл. 3. У более толстых пленок имеется узкий провал в спектре вблизи ϵ_F . Таким образом, на ультратонких пленках, где малая толщина является главным параметром эффективного разупорядочения и перехода Андерсона, тоже есть и минимум в спектре на металлической стороне перехода, и кулоновская щель со стороны изолятора. Главная особенность и отличие кривых на рис. Б.6 в том, что они демонстрируют спектры заведомо двумерных систем.

Рис. Б.1 и формулы (Б.1)–(Б.5) описывают простейшую схему туннельного эксперимента при практически неограниченных размерах электродов и площади туннельного контакта. В этой схеме все особенности вольт-амперной характеристики обусловлены плотностью состояний массивного материала. Представим себе теперь, что один из электродов — это маленькая металлическая гранула, все размеры которой малы; пусть это шар радиуса a . Если на эту гранулу протуннелирует один электрон, то гранула станет заряженной и вокруг нее возникнет

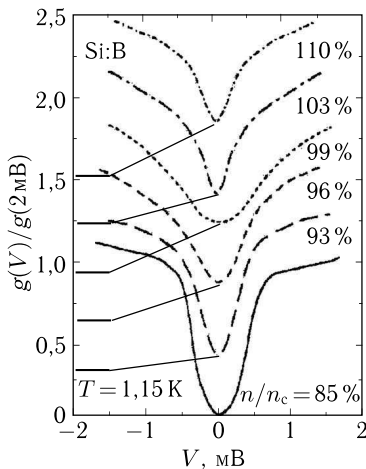


Рис. Б.5. Дифференциальная проводимость туннельных контактов с Si:B в качестве электрода после исключения влияния теплового уширения. Кривые смещены для ясности, но для каждой кривой на оси ординат указан соответствующий ей ноль [7]

электрическое поле. Энергия поля уединенного металлического шара с зарядом e равна

$$U = \frac{e^2}{2C}, \quad C = \kappa a, \quad (\text{Б.6})$$

где C — это емкость уединенного шара, а κ — это диэлектрическая проницаемость окружающего его изолятора. Если температура $T \ll U$, то туннелирование станет возможным только, когда напряжение между шаром и массивным электродом превысит U/e : $V = e/C = e/\kappa a > U/e$. Этот порог по напряжению называется кулоновской блокадой (см. также (8.6)–(8.11), гл. 8).

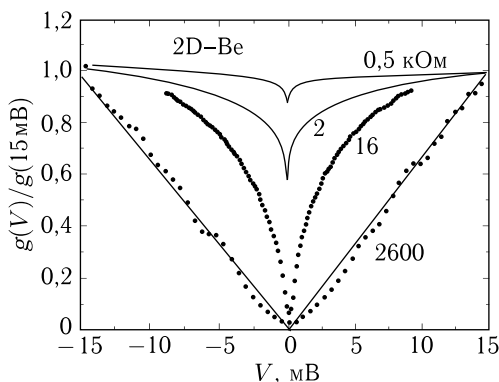


Рис. Б.6. Туннельные проводимости при температуре $T = 50$ мК контактов с пленками Be с различными сопротивлениями. Из-за слишком большого сопротивления при $T = 50$ мК измерения на самой высокоомной пленке были проведены при $T = 700$ мК [8]

Для туннельного эксперимента нужны контакты к обоим электродам. Чтобы такой контакт не сделал маленький металлический электрод большим, его можно расположить между двумя массивными электродами, устроив два туннельных промежутка последовательно. На рис. Б.7 демонстрируются принципиальная схема и результаты такого эксперимента. На слегка окисленную поверхность алюминиевой пленки напыляли олово, которое собиралось в островки; средний размер островков зависел от количества напыленного олова. Затем образец подвергался дальнейшему окислению, в результате чего между частицами олова образовывался толстый слой окиси алюминия, а сами частицы покрывались тонким слоем окиси олова. Затем сверху напыляли слой алюминия для создания верхнего электрода (см. вставку на рис. Б.7).

Кривые, представленные на рис. Б.7, были получены в магнитном поле, достаточно большом для того, чтобы разрушить сверхпроводимость. Тогда все электроды заведомо являются хорошими металлами, и вольт-амперные характеристики структуры должны были бы быть

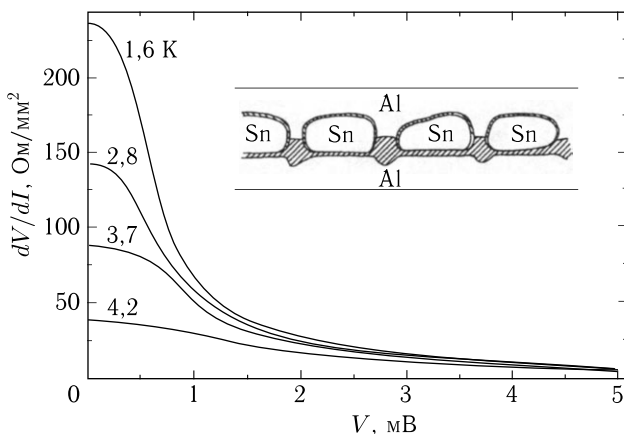


Рис. Б.7. Туннельные характеристики структуры Al — гранулы Sn — Al при различных температурах. Схема структуры показана на вставке, [9]

горизонтальными прямыми. Падение проводимости, наблюдаемое при малых напряжениях и нарастающее при понижении температуры, есть следствие малости промежуточных электродов — гранул олова, т.е. следствие кулоновской блокады.

Кулоновская блокада является широко исследуемым и используемым явлением при изучении наноструктур. В них обычно промежуточным электродом является одиночная гранула. Мы здесь сталкиваемся с кулоновской блокадой в другом предельном случае, когда гранул очень много. В эксперименте, представленном на рис. Б.7, гранулы образуют двумерный слой. Совокупность гранул может образовывать и трехмерный объемный конгломерат. Он называется гранулированным металлом, и ему посвящена отдельная глава этой книги (гл. 8). Там, в частности, описаны и измерения эффективной плотности состояний в гранулированном металле, выполненные при помощи туннельной методики. При этом гранулированный металл являлся не промежуточным электродом, как на рис. Б.7, а одним из двух основных электродов, в соответствии со схемой рис. Б.1.

Имея в виду дальнейшее развитие туннельной методики и ее возможностей, следует обратить внимание на то, что формула (Б.1) предполагает, что при туннелировании не сохраняется волновой вектор. Несохранение нормальной по отношению к поверхности контакта компоненты волнового вектора естественно; оно обусловлено нарушением однородности пространства в соответствующем направлении. Тангенциальная компонента не сохраняется лишь из-за шероховатости поверхности контакта. Если барьер контакта изготовить с атомно гладкими краями, то можно рассчитывать на туннелирование с сохранением тангенциальной компоненты волнового вектора. Такое туннелирование следует назвать когерентным.

Когерентное туннелирование было реализовано в специальных экспериментах [10], схема и результаты которых приведены на рис. Б.8. Туннелирование осуществлялось через тонкий барьер между двумя двумерными квантовыми ямами. Ямы из двух слоев GaAs толщиной $140 \text{ \AA}$ каждый с барьером из AlAs толщиной $70 \text{ \AA}$ между ними были изготовлены методом молекулярной эпитаксии. С обеих сторон эта структура была прикрыта слоями $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$. Общая толщина этого пятислойного сэндвича была около 50 мкм . С двух торцов пластины были вожжены индиевые контакты 1 и 2, через каждый из которых можно было осуществлять электрический контакт с обеими ямами с двумерным электронным газом, а сверху и снизу были напылены металлические затворы (см. вставку на рис. Б.8). Затворы a_1 и a_2 использовались для того, чтобы разрезать двумерный газ в соответствующей яме на две части и превратить параллельное подсоединение двумерных ям к индиевым контактам в последовательное, с включением между ними туннельного промежутка (затвор a_1 отсекает двумерный газ в нижней яме от контакта 2, а затвор a_2 — двумерный газ в верхней яме от контакта 1). При помощи затворов t_1 и t_2 можно было менять концентрацию носителей, соответственно, в нижней и верхней ямах.

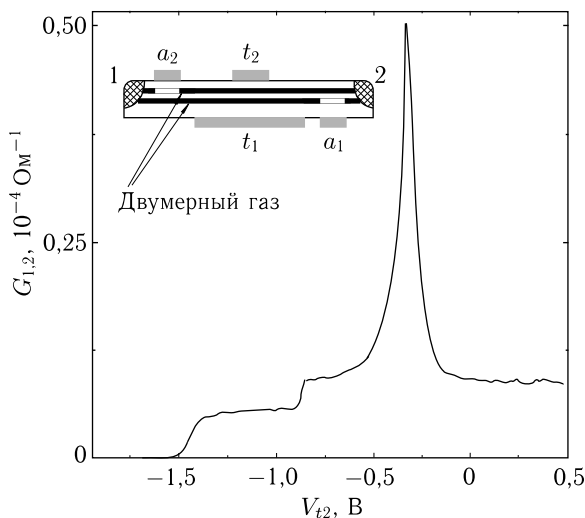


Рис. Б.8. Зависимость проводимости G_{12} между контактами 1 и 2 и пропорционального ей туннельного тока J от напряжения на затворе t_2 при фиксированном напряжении на затворе t_1 . При этом, как показано на вставке, нижний электронный газ отсечен запирающим напряжением на затворе a_1 от контакта 2, а верхний — напряжением на затворе a_2 от контакта 1 [10].

Температура $T = 1,5 \text{ К}$, тянущее напряжение $v = 0,1 \text{ мВ}$

При выращивании этой структуры пришлось решать две сложные экспериментальные задачи, потребовавшие жесткого контроля процесса молекулярной эпитаксии. Во-первых, требовалось получить атомно гладкую границу между слоями GaAs и AlAs на очень большой площади. Во-вторых, электронные плотности в двух ямах должны были изначально примерно одинаковы.

Посмотрим, как должен вести себя туннельный ток при выполнении обоих законов сохранения, и энергии и импульса. Соединив контакты 1 и 2 внешней цепью, в которой имеется только пренебрежимо маленькая измерительная разность потенциалов v , мы сравниваем ферми-уровни по обе стороны барьера и автоматически получаем выполнение закона сохранения энергии. В двумерном газе весь волновой вектор лежит в плоскости и равен

$$k_F = (2\pi n)^{1/2}. \quad (\text{Б.7})$$

Поэтому в нулевом магнитном поле когерентное туннелирование возможно только при одинаковых концентрациях электронов в ямах.

Приведенная на рис. Б.8 кривая получена при тянущем напряжении $v = 0,1$ мВ и при некотором фиксированном напряжении на нижнем затворе V_{t1} . Она демонстрирует зависимость туннельного тока $J = vG_{1,2}$ от напряжения на верхнем затворе V_{t2} , т.е. от концентрации электронов n_2 в верхней яме. Пик туннельного тока возникает при равенстве концентраций в обеих ямах, $n_1 = n_2$. Как и должно быть, он смещается при изменении V_{t1} . Относительная величина пика определяет отношение когерентного и некогерентного туннелирования.

Картину дополняют две ступени на кривой $J(V_{t2})$ слева от когерентного пика. Первая ступень возникает тогда, когда концентрация электронов в верхней яме под затвором t_2 уменьшается настолько, что под ним получается изолятор. Тогда площадь туннельного контакта уменьшается примерно вдвое. Вторая ступень возникает при том напряжении V_{t2} , когда электрическое поле от затвора t_2 разрезает область, занятую электронами в нижней яме, и полностью разрывает электрическую цепь между контактами 1 и 2.

Б.1. Сканирующая туннельная спектроскопия¹⁾

Во всех рассуждениях выше предполагалось, что площадь S туннельного контакта велика на масштабах межатомных расстояний:

$$S \gg a^2, \quad (\text{Б.8})$$

а все обсуждавшиеся до сих пор эксперименты были сделаны именно на таких контактах. Между тем, в последние десятилетия появилась и бурно развилась новая область экспериментальной физики —

¹⁾ Подробное последовательное изложение физики и техники туннельной сканирующей спектроскопии можно найти в [11].

сканирующая туннельная спектроскопия, которая все чаще успешно применяется к задачам об электронном спектре в твердых телах.

В сканирующем туннельном микроскопе одним из электродов является металлическая игла с радиусом закругления на конце порядка нескольких ангстрем, а туннелирование происходит через вакуумный промежуток между концом иглы и плоским вторым электродом. Работоспособность такой конфигурации обеспечивается возможностью тонкого и точного контролируемого перемещения иглы относительно исследуемой поверхности проводящего материала, которая служит вторым электродом. Платформа с иглой прикреплена к конструкции из пьезоэлектрической керамики. Деформируясь под действием прикладываемых электрических напряжений, эта конструкция перемещает иглу перпендикулярно исследуемой поверхности и в двух направлениях вдоль нее. Наличие обратной связи по туннельному току позволяет поддерживать туннельный промежуток постоянным.

Новая экспериментальная технология привела к изменению терминологии. Исследуемым объектом в сканирующем микроскопе является один берег туннельного контакта, а игла служит для собирания туннельного тока. Входящие в экспоненту в коэффициенте A в уравнении (Б.1) величины m_{ef} и U_{ef} относятся именно к исследуемому объекту. Часто в качестве m_{ef} подставляют массу свободного электрона, так что свойства объекта описываются величиной U_{ef} , которая в этом случае называется работой выхода и обычно обозначается буквой W . Смысл этого названия подчеркивается тем обстоятельством, что ширина туннельного промежутка w является теперь экспериментально контролируемой переменной величиной, а вероятность q найти электрон на расстоянии w от поверхности исследуемого объекта равна

$$q(w) = |\psi_n(0)|^2 e^{-2\kappa w}, \quad \kappa = \hbar^{-1} \sqrt{2mW}. \quad (\text{Б.9})$$

Типичная величина работы выхода $W \sim 5$ эВ, а κ обычно равна по порядку величины $1 \text{ \AA}^{-1}$.

Потребовалось также уточнение определения функции плотности состояний $g(\varepsilon)$, которая входит в формулы (Б.1)–(П.Б.5). На протяжении всей книги под плотностью состояний g понималась величина, усредненная по достаточно большому объему. На самом деле и электронная плотность n и ее производная $g = \partial n / \partial \varepsilon$ являются функциями координат: $n = n(\mathbf{r})$, $g = g(\varepsilon, \mathbf{r})$. Самые мелкие осцилляции этих функций определяются структурой волновых функций и имеют атомный масштаб. Чтобы отличать функцию $g(\varepsilon, \mathbf{r})$ от усредненной функции плотности состояний, ее называют локальной плотностью состояний (LDOS). Формула (Б.9) по существу описывает уменьшение LDOS по мере удаления от поверхности исследуемого объекта. Фактически обратная связь по туннельному току стремится удерживать кончик иглы на поверхности $g(\mathbf{r}) = \text{const}$.

Естественно было ожидать, что разрешение сканирующего туннельного микроскопа будет определяться диаметром иглы и составлять несколько межатомных расстояний. На самом деле оно оказалось выше и позволяло увидеть отдельные атомы. Причин «такого везения» несколько. Мы здесь упомянем только одну из них.

Представим поверхность кончика иглы в виде полусферы радиуса R и пусть расстояние от полюса сферы до ближайшей точки $x = 0$ на исследуемой плоской поверхности равно w_0 (см. рис. Б.9). При малых x расстояние от точки x на плоской поверхности до иглы на $\Delta w = x^2/2R$ больше, чем в центре. Согласно формуле (Б.1), это означает, что ток через площадку dS на расстоянии x от оси иглы в

$$i(x)/i(0) = \exp(-2\kappa x^2/2R) \quad (\text{Б.10})$$

раз меньше, чем через площадку той же площади непосредственно на оси. Полагая $\kappa \approx 1 \text{ \AA}^{-1}$, получим, что под иглой радиусом $R = 20 \text{ \AA}$ плотность туннельного тока падает в 10 раз на расстоянии всего $6 \text{ \AA}$ от оси.

Благодаря возможности осуществлять спектроскопию плотности состояний на малом участке поверхности, т.е. измерять зависимость функции LDOS $g(\varepsilon, \mathbf{r})$ от обеих переменных, сканирующий электронный микроскоп открывает новые области экспериментальных исследований электронных систем, в частности, квантовых фазовых переходов. В качестве примера на рис. Б.10 приведены измерения LDOS при разных энергиях в пределах одного, наинизшего, уровня Ландау в двумерном электроном газе в условиях целочисленного квантового эффекта Холла [12]. Уровень уширен и имеет вид миниполосы из-за пространственных флуктуаций случайного потенциала. Два листа пространственного распределения функции $g(\varepsilon, \mathbf{r})$ с наименьшим и наибольшим значениями V_s соответствуют крыльям уширенного уровня, а листы со средними значениями V_s — его центральной части.

Заметьте. LDOS, регистрируемая в описываемых экспериментах, формируется только за счет делокализованных электронов. Туннелирование между иглой и локализованным состоянием в первом приближении не отражается на токе, регистрируемом через контакты. На листе $V_s = -116,3 \text{ мВ}$ видны две дефектные области, гальванически связанные с контактами. Наличие этих дефектов сказывается и на противоположном крыле уровня Ландау при $V_s = -89,0 \text{ мВ}$.

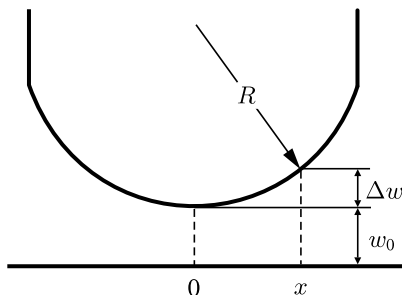


Рис. Б.9. Схематическое изображение иглы туннельного микроскопа

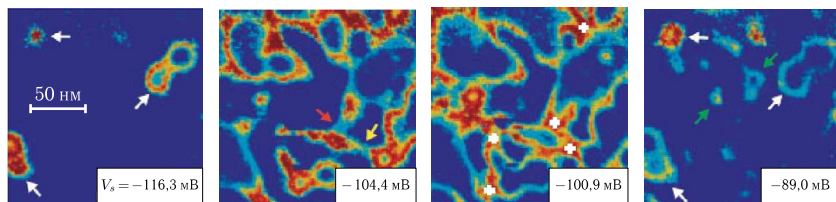


Рис. Б.10. LDOS в пределах низшего уровня Ландау в двумерном электронном газе на поверхности $n\text{-InO}(110)$, легированной 0,01 монослоя Cs, в поле 12 Тл при температуре 0,3 К. На каждом листе указано напряжение на игле V_s ($V_s = 0$ соответствует уровню Ферми). Значения $V_s = -116,3$ мВ и $V_s = -89,0$ мВ соответствуют крыльям минизоны Ландау, а значения $V_s = -104,4$ мВ и $V_s = -100,9$ мВ — ее центральной части. Значения dI/dV растут от темных к светлым [12]

Рис. Б.10 иллюстрирует важные особенности поведения электронов на уровнях Ландау в условиях квантового эффекта Холла, которые обсуждались в гл. 9: наличие замкнутых спиральных траекторий вдоль эквипотенциалей вокруг экстремумов случайного потенциала и перколяционной сетки делокализованных состояний в середине миниполосы уровня Ландау.

Б.2. Контрольные вопросы и задачи

1. Волновая функция электрона затухает с расстоянием x от поверхности металла как

$$\psi(x) = \psi(0)e^{\kappa x}, \quad \kappa = \hbar^{-1}\sqrt{2mW},$$

где работа выхода W — разность между энергией электрона в вакууме и на уровне Ферми в металле. Полагая $W \sim 5$ эВ, оценить, во сколько раз уменьшится туннельный ток при увеличении туннельного промежутка на $1 \text{ \AA}$.

Примерно в e раз.

2. Одна из поверхностей туннельного контакта — плоскость, а другая — сфера радиусом $50 \text{ \AA}$. Работа выхода W . На какой площади под сферой плотность туннельного тока отличается от плотности тока в центре менее, чем в 10 раз? Имеет ли значение минимальное расстояние от плоскости до сферы?

$$\pi r^2, \quad r = 10, 8 \text{ \AA}.$$

Список литературы

1. Туннельные явления в твердых телах. — М.: Мир, 1973 (перевод книги Tunneling phenomena in solids, eds. E. Burstein, S. Lundquist. — Plenum, 1969).

2. *Солимар Л.* Туннельный эффект в сверхпроводниках и его применение. — М.: Мир, 1974 (перевод книги *Solymar L.* Superconductive tunneling and applications. — Chapman and Hall, 1972).
3. *Aleiner I. L., Brouwer P. W., Glazman L. I.* Quantum effects in Coulomb blockade // *Phys. Rep.* **358**, 309 (2002).
4. *Massey J. G., Lee M.* // *Phys. Rev. Lett.* **75**, 4266 (1995).
5. *McMillan W. L. Mochel J.* // *Phys. Rev. Lett.* **46**, 556 (1981).
6. *Hertel G., Bishop B. J., Spencer E. G., Rowell J. M. Dynes R. C.* // *Phys. Rev. Lett.* **50**, 743 (1983).
7. *Massey J. G., Lee M.* // *Phys. Rev. Lett.* **77**, 3399 (1996).
8. *Butko V. Yu., DiTusa J. F., Adams P. W.* // *Phys. Rev. Lett.* **84**, 1543 (2000).
9. *Giaever I., Zeller H. R.* // *Phys. Rev. Lett.* **20**, 1504 (1968).
10. *Eisenstein J. P., Pfeiffer L. N., West K. W.* // *Appl. Phys. Lett.* **58**, 1497 (1991); *Phys. Rev. B* **44**, 6511 (1991)
11. *Chen C. J.* Introduction to Scanning Tunneling Microscopy. Oxford University Press, 1993.
12. *Hashimoto K., Sohrmann C., Wiebe J., Inaoka T., Hirayama Y., Römer R. A., Wiesendanger R., and Morgenstern M.* // *Phys. Rev. Lett.* **101**, 256802 (2008).

УКАЗАТЕЛЬ МАТЕРИАЛОВ ¹⁾

- Ag, пленка 6.10
Al, пленка 2.10
Al, гранулир. 8.3
Al–Pd–Mn, квазикристал. 7.10
Al–Pd–Re, квазикристал. 7.11, 7.12
Au, пленка 2.3, 6.10
Au, гранулир. 8.2, 8.5, 8.6
Be, пленка 3.10, 4.7, Б.6
Bi, аморф., пленка 6.16
Cs–Au 7.6, 7.7
Cu, пленка 2.3, 2.13, 6.10
Cu–Ag 5.10
GaAs 4.5, 6.11, 6.14, 9.26
GaAs/GaAlAs, гетероструктура
2.16, 5.11, 6.9, 9.7–9.9, 9.18–9.21,
9.23, 9.24,
Ga_{0.15}Al_{0.85}As–Ga_{0.67}Al_{0.33}As 10.7,
10.8,
Ga_xMn_{1–x}As 6.6, 6.7
p-Ge 3.6, 4.1
GeAs–AlAs–GeAs Б.8
Ge–Au Б.3
Ge:As 6.4, 6.5
Ge:Ga 4.2
Ge:P, Ge:Sb 3.5
Ge/SiGe, гетероструктура 9.27
In, гранулир. 8.1
InGaAs/InP, гетероструктура 2.15
InO, аморф. 2.4
InO, поверхность (110), легирован-
ная Cs Б.10
InP 4.4
Li, пленка 2.10
Mg, пленка 2.9, 2.10, 2.14
Nb₃Sb, Nb₃Sn 1.4
Ni, гранулир. 8.5–8.7
Pb 8.4
Sb 2.11, 2.12
Sb–Cd, гранулир. 8.8, 8.9
Sb–Ga, гранулир. 8.8, 8.10
Si:As 4.6
Si:B 3.9, 4.2, 4.8–4.10, Б.2, Б.5
Si–Nb Б.4
Si:P 5.9
Si, MOS-структура 2.5, 5.14, 5.15,
6.12, 9.5, 9.6, 9.17
TiAl 1.5
WO₂ 1.7
Y 1.6
Щелочной металл–Pb 7.2–7.5
Щелочной металл–Sn 7.4, 7.5

¹⁾ Ссылки приведены на номера рисунков, а не страниц.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ ¹⁾

Абрахамса–Миллера сетка 76

Атомные конфигурации

— — Зинтля 160

— — локальные в квазикристаллах 164

Виртуальные связанные состояния 170

Всплывание уровней 224, 232

Время

— магнитное, τ_B 36

— мнимое 241

— рассасывания электронной неоднородности 276

— рассеяния упругого, τ 27

— — электрон-фононного, τ_{ph} 23, 54

— — электрон-электронное, баллистическое, τ_e^{ball} 49

— — электрон-электронное, диффузионное, τ_e^{diff} 50

— расфазировки, τ_{ee} 49

— сбой фазы, τ_φ 27

— спин-орбитальное, τ_{so} 42

— эффективное, τ_{eff} , 23

Грюнайзена функция 19, 24

Диаграмма

— критической области квантового фазового перехода 245

— — — перехода металл–изолятор 130, 249

— однопараметрического скейлинга 125, 151,

— окрестности виртуального перехода 2D-газа 250

— потока 229, 232

— прыжковой проводимости 83

— совместного размерного и магнитного квантования 198

— частота столкновений–температура 24

— электронная концентрация–беспорядок 101

Димерная модель 115

Диффузия электронов 28, 50, 131

Длина

— волны электрона, $1/k_F$ 11, 102

— диффузионная, L_φ , L_{ee} 30, 49

— магнитная, см. Радиус магнитный

— корреляционная, ξ 123, 126, 189, 265, 289

— локализации, ξ 74, 141, 179, 266

— свободного пробега, l 11, 16, 111

Интеграл перекрытия 92, 99, 100

Инвариант в задачах связей, I_b 260, 261

— в задачах узлов, I_s 260, 261, 264

Кермет 176

Класс универсальности 129

Компьютерное моделирование 68, 70, 97, 118, 260, 262, 268

Корбино диск 202, 206, 213, 218

Корреляции

— в случайном потенциале 115, 118

— актов рассеяния 21, 115

Край подвижности 94, 118

Локализационный параметр, $a_B n^{-1/3}$ 98, 100

¹⁾ В указатель не включены термины и понятия, фигурирующие в оглавлении.

Мезоскопия 41

Металл

- 2D при $T = 0$ 145, 154
- при $T = 0$, определение 90
- жидкий 13, 157
- стандартный 12, 18, 157

Микроволновый бильярд 119

Мягкая локализация 196, 200

Мягкая щель 69, 88, 185, 273-278

Нелокальность проводимости 40, 106

Обход точки фазового перехода 133

Параллельные каналы проводимости

- в металлах с сильным рассеянием 20
- в полупроводниках 78

Плотность состояний 51

- — локальная 282
- — туннельная 276

Покрытие Пенроуза 164, 165

Работа выхода 282

Радиус

- локализации, *см.* Длина локализации
- магнитный, r_B 36, 196
- перколяционный, r_c 78, 263
- экранирования, r_e 97
- электронной орбиты на N-том уровне Ландау, r_N 199

Размерность

- $2 + \varepsilon$ 126, 129
- в перколяционных задачах 261, 263
- , влияние на квантовые поправки к проводимости 30, 51, 52
- , влияние на переход металл-изолятор 105

Распределение log-нормальное 139

Рассеяние назад в краевых каналах 214

- — в одномерных системах 108
- — в условиях слабой локализации 34

Резервуар, см. Термостат

Резонансные ямы 93, 95

Сверхпроводимость 153, 178, 179, 185, 189, 273

Случайное электрическое поле 61

Случайный потенциал короткодействующий 201, 214

— — крупномасштабный 200, 210

Статический скин-эффект 196, 215

Статистическая сумма 237

Структурный фактор 14

Термостат (или резервуар) 44, 106, 208

Токовые каналы

- — в одно- и многоканальных 1D-проводниках 105
- — во фрактально-гранулированной среде 189
- — идеальные 105, 214
- — краевые 213
- —, сетка в перколяционной системе 269

Универсальные флуктуации кондактанса 39

Флуктуации квантовые 240

- мультифрактальной функции 136

Фрактальная размерность 137

Холловский изолятор 222, 230

- квантовые жидкости 221, 230
- проводимость 195, 202
- сопротивление 195, 202
- ток 215

Цилиндрические пленки 38

Частота столкновений эффективная 23

— циклотронная 36, 194

Шунтирующее сопротивление 20

Эйнштейна соотношение 132