

ПОВЕРХНОСТНЫЕ
(ТАММОВСКИЕ)
СОСТОЯНИЯ

С. Дэвисон, Дж. Левин

ИЗДАТЕЛЬСТВО «МИР»

SOLID STATE PHYSICS
ADVANCES IN RESEARCH AND APPLICATIONS
EDS. H. EHRENREICH, F. SEITZ, D. TURNBULL
VOLUME 25

SURFACE STATES

by

S. G. DAVISON

DEPARTMENT OF PHYSICS,
INSTITUTE OF COLLOID AND SURFACE SCIENCE,
CLARKSON COLLEGE OF TECHNOLOGY,
Potsdam, New York

and

J. D. LEVINE

RCA LABORATORIES,
DAVID SARNOFF RESEARCH CENTER,
Princeton, New Jersey

ACADEMIC PRESS
NEW YORK AND LONDON
1970

С. ДЭВИСОН, ДЖ. ЛЕВИН

**ПОВЕРХНОСТНЫЕ
(ТАММОВСКИЕ)
СОСТОЯНИЯ**

Перевод с английского А. Я. БЕЛЕНЬКОГО

Под редакцией
д-ра физ.-мат. наук Д. А. КИРЖНИЦА

ИЗДАТЕЛЬСТВО «МИР»

МОСКВА 1973

Еще в 1932 г. выдающийся советский ученый И. Е. Тамм указал на возможность существования электронных поверхностных уровней (получивших позднее название «таммовских уровней»), которые играют важную роль в физике твердого тела. Однако до сих пор квантовая теория кристалла, ограниченного поверхностью (в отличие от теории бесконечного кристалла), развита недостаточно. Несмотря на наличие многочисленных оригинальных исследований, достаточно полные обзорные работы в этой области практически отсутствуют.

Данный монографический обзор С. Дэвисона и Дж. Левина (США) восполняет этот пробел. Авторы обобщили и систематизировали обширный материал, накопленный за 40 лет (особенно за последние годы).

Книга представляет большой интерес для специалистов в области физики твердого тела.

© Перевод на русский язык, «Мир», 1973

Редакция литературы по физике

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Более сорока лет назад в 1932 г. появилась статья Игоря Евгеньевича Тамма «О возможной связи электронов на поверхности кристалла»¹⁾. В этой статье было показано, что граница идеальной кристаллической решетки может сама по себе послужить источником особых состояний электрона, локализованных вблизи этой границы. Такие поверхностные состояния, названные впоследствии «таммовскими» (иногда говорят также об «уровнях Тамма»), отщепляются от разрешенной области спектра и располагаются внутри запрещенной зоны. По своей природе они во многом похожи на обычные связанные состояния, изучаемые в курсах квантовой механики. И те и другие характеризуются энергией, лежащей в недоступной для «свободной» частицы области спектра. И те и другие описываются экспоненциально затухающей волновой функцией. И те и другие отвечают полюсам матрицы рассеяния (для таммовских состояний — рассеяния блоховской волны поверхностью кристалла).

С возникновением понятия о поверхностных состояниях стало ясно, что поверхность кристалла играет роль самостоятельной его подсистемы, причем принадлежащие ей электроны также движутся в периодическом (двумерном) поле. Стало возможным говорить о таких смешанных структурах, как металл с диэлектрической поверхностью или, напротив, диэлектрик, на поверхности которого расположен двумерный металл. Обсуждалась даже модель кристалла со сверхпроводящей поверхностью. Впрочем, создать кристалл с поверхностью, сколько-нибудь близкой к идеальной, исключительно трудно. Поэтому упомянутые экзотические структуры до сих пор непосредственно не наблюдались. Несмотря на это, таммовским и поверхностным состояниям иного типа принадлежит огромная роль во всей три совокупности явлений, которые называются поверхностными.

Поверхностным состояниям посвящены сотни работ. Однако до последнего времени отсутствовали достаточно полные обзоры, охватывающие всю проблему поверхностных состояний в целом. В особенности это относится к литературе на русском языке, где имеется только небольшой обзор И. М. Лифшица и С. И. Пекара²⁾, написанный около 20 лет назад и охватывающий лишь теоретические аспекты проблемы.

Можно надеяться, что этот пробел будет восполнен изданием настоящей книги, которая представляет собой перевод большого обзора из 25-го тома *Solid State Physics* и содержит подробное изложение обширного круга теоретических и экспериментальных вопросов, относящихся к физике поверхностных состояний. Авторы книги, американские физики С. Дэвисон и Дж. Левин, — известные специалисты

¹⁾ И. Е. Тамм, *Phys. Zs. Sowjet.*, 1, 733 (1932).

²⁾ И. М. Лифшиц, С. И. Пекар, *УФН*, 56, 531 (1955).

в области физики поверхностных явлений, внесшие в нее немалый вклад. К русскому изданию авторы любезно прислали предисловие, а также ряд дополнений к тексту и литературе.

О содержании книги лучше всего говорит ее оглавление. При переводе текст остался практически неизменным, исключая расшифровку многочисленных сокращений. Краткие примечания переводчика и добавленные им ссылки на источники давались лишь в тех случаях, когда это казалось явно необходимым.

Книга свидетельствует о бурном развитии одной из ранних идей замечательного нашего физика И. Е. Тамма. Она непосредственно адресуется специалистам в области физики твердого тела — экспериментаторам и теоретикам, а также аспирантам и студентам соответствующего профиля, которым она принесет несомненную пользу.

Д. Киржниц

ПРЕДИСЛОВИЕ АВТОРОВ К РУССКОМУ ИЗДАНИЮ

В 1932 г. И. Е. Тамм в своей теоретической работе показал, что электроны могут быть связаны с поверхностью твердого тела, перемещаясь только вдоль нее. Эти электронные поверхностные состояния, известные с тех пор как *таммовские состояния*, играют ведущую роль в механизмах, лежащих в основе работы ряда твердотельных устройств. Свидетельством истинной значимости работы И. Тамма служит огромное число ссылок на нее в обширной литературе по поверхностным состояниям, накопленной за прошедшие годы.

В настоящем обзоре сделана попытка в сжатой форме, но с достаточной полнотой, охватить как теорию, так и экспериментальные данные. Можно надеяться, что эта книга окажется полезной для физиков, химиков и инженеров по радиоэлектронике. Книга может быть также использована студентами в качестве пособия при изучении общих курсов квантовой механики и физики твердого тела.

Авторы приносят свои благодарности Т. Фишеру, С. Фриману, А. Гудману, П. Марку, М. Стенслицкой, Р. Сванку и К. Вронскому за их многочисленные ценные замечания и полезные предложения. Мы также благодарны А. Граер за ее терпение, проявленное при первоначальной перепечатке рукописи, и П. Дэвисон за отличную окончательную перепечатку и помощь в чтении корректуры.

Наконец, мы чрезвычайно признательны Д. Киржницу, любезно согласившемуся отредактировать перевод, и А. Беленькому за тщательный перевод книги.

Мы считаем большой честью для себя, что наша работа признана достойной воздать должное памяти такого выдающегося ученого, как лауреат Нобелевской премии И. Е. Тамм.

Март 1973 г.

С. Дэвисон,
Дж. Левин

Глава I

ВВЕДЕНИЕ

В имеющихся обзорах [1—5]¹⁾, посвященных квантовомеханическим аспектам теории поверхностных явлений, содержится преимущественно обсуждение отдельных частных вопросов или методов. В данной работе подобное ограничение отсутствует: рассмотрены большинство проблем и методов, которые интересуют исследователя, занимающегося поверхностными явлениями. Это достигается охватом широкого круга вопросов и их анализом с различных точек зрения. Обзор содержит как теоретическое описание, так и экспериментальную характеристику поверхностей кристалла, причем особое внимание уделено собственным поверхностным состояниям. Можно, таким образом, надеяться, что получена весьма полная картина, отражающая современный уровень понимания свойств электронов у чистых поверхностей твердых тел. Дополнительный теоретический материал читатель сможет найти в работе Дэвисона и Стенслицкой [6], а экспериментальный — в работах Мейни с соавторами [7] и Франкля [8]. Сюда же следует отнести книгу Волькенштейна [9]. Она рекомендуется тем, кто интересуется проблемами катализа, которые оказались в основном вне поля зрения настоящего обзора. Последний может быть также интересен и для специалистов, работающих в других областях (физика спиновых волн, фононов, экситонов и т. д.), в которых проблемы поверхности играют важную роль.

Поскольку кристалл представляет собой систему многих частиц, содержащую движущиеся ядра и электроны, состояния этих частиц описываются *полным* многочастичным уравнением Шредингера. Пользуясь приближением Борна—Оппенгеймера [10], можно выделить из этого уравнения ядерную и электронную части [11]. *Электронное* уравнение Шредингера, которое мы будем

¹⁾ См. также [317]. — *Прим. перев.*

здесь рассматривать, представляет собой многоэлектронное волновое уравнение, описывающее движение электронов в поле фиксированных ядер¹⁾. Используя модель независимых частиц в рамках метода Хартри — Фока [11, 12], т. е. предполагая, что каждый электрон испытывает действие статического потенциала ядер и *среднего* поля всех других электронов, можно свести многоэлектронное уравнение Шредингера к одноэлектронному виду. Таким образом, анализ электронных свойств поверхности кристалла сводится к решению одноэлектронного уравнения Шредингера совместно с граничными условиями, отвечающими рассматриваемой системе. Методы, которые при этом используются, будут обсуждены в последующих разделах.

¹⁾ Применимость этого так называемого «адиабатического приближения» в задаче о поверхностных состояниях далеко не очевидна. Характерный интервал энергий электронных состояний составляет примерно 0,1 эВ, что может быть сравнимо с характерными фоновыми энергиями (10^{-2} эВ). Ярче всего отсутствие «адиабатичности» проявляется в существовании «перестройки» поверхности (см. ниже) — явления, в котором электроны играют важную роль. — *Прим. перев.*

Глава 2

КРАТКИЙ ИСТОРИЧЕСКИЙ ОЧЕРК

Одно из самых первых приложений квантовой теории к задаче об идеальном кристалле было основано на использовании предложенной в 1931 г. модели Кронига — Пенни [13] (линейная цепочка δ -образных потенциалов отталкивания). Годом позже И. Е. Тамм [14] рассмотрел кристалл с таким потенциалом, ограниченный свободной поверхностью. Поверхность была введена как ступенчатый скачок потенциала, причем волновая функция и ее производная «сшивались» в точке разрыва. Тамм был первым, кто показал, что при определенных условиях могут появиться поверхностные состояния. Затем Мауэ [15] рассмотрел модель кристалла в приближении почти свободных электронов, разлагая потенциал в ряд Фурье. И в этой работе поверхность вводилась как ступенчатый скачок периодического потенциала в точке его максимума или минимума и были найдены условия существования поверхностных состояний. В 1939 г. Гудвин опубликовал три работы [16], посвященные поверхностным состояниям. В первой он обобщил результаты Мауэ, но две другие содержали подход, бывший для того времени новым. В основе этого подхода лежал метод линейных комбинаций атомных орбиталей (ЛКАО). Существенно, что при этом разрывы физических величин на поверхности описывались иным способом: предполагалось просто, что поверхностным атомам отвечают значения кулоновских и резонансных интегралов, отличные от соответствующих значений для атомов в объеме. Гудвин рассмотрел кристаллы, обладающие только одной орбиталью на атом, а также кристаллы с *sp*-гибридизацией. В том же 1939 г. Шокли [17] изучал свойства одномерного периодического потенциала общего вида, ограниченного в точке максимума ступенькой. В его методе также использовалось сшивание волновых функций на разрыве. Шокли показал, что если зоны объемных состояний «перекрываются», то поверхностные состояния

возникают при условии, что поверхностное «возмущение» достаточно мало. Это противоречило предыдущим выводам, согласно которым поверхностные состояния возникают, только если поверхностное возмущение достаточно велико. Позднее эти два вида поверхностных состояний были названы соответственно «состояниями типа Шокли» и «таммовскими состояниями». Шокли предположил, что перекрывающиеся зоны и пренебрежимо малое поверхностное возмущение характерны для ситуации в кристаллах Si и Ge.

Итак, в 40-х годах теория поверхностных состояний была уже до некоторой степени разработана. Напротив, экспериментальные доказательства их существования стали поступать только с конца 40-х годов с развитием полупроводниковой техники и технологии очистки кристаллов. В 1947 г. Бардин [18] объяснил «аномальные» результаты Мейергофа [19], полученные для контактов металла с кремнием, на языке поверхностных состояний. В 1948 г. Шокли и Пирсон [20] проверили предположение Бардина. Накладывая однородное поле, они изменяли потенциал свободной поверхности полупроводника (так называемый эксперимент с «эффектом поля»). Оказалось, что при наложении дополнительного внешнего поля параллельно поверхности часть поверхностного заряда начинает двигаться в то время, как большая его доля остается неподвижной. Эта неподвижная часть заряда была приписана электронам, захваченным в поверхностные состояния. Позднее было найдено, что существуют как «медленные», так и «быстрые» ловушки; они связаны с образованием окисла на поверхностях Si и Ge под действием воздуха.

Эти экспериментальные результаты вновь усилили в 50-х годах интерес к теории поверхностных состояний. Особенно это относится к опытам с поверхностями Si и Ge. Статцем [21], Артманом [22], Гоффманом [23], Липпманом [24] и другими был разработан ряд вопросов, связанных с состояниями Тамма и Шокли. Следующий шаг в направлении создания подробной общей теории локализованных поверхностных состояний, приложимой к реальным системам, был сделан в 1957 г. Коутецким [25] и в 1960 г. Коутецким и Томашеком [26]. Этим авто-

рам удалось рассмотреть все типы состояний, локализованных у поверхности идеального кристалла. В их методе, названном методом одноэлектронной резольвенты, большинство результатов было получено в представлении ЛКАО. В конце этого периода Антончик [27], видоизменив подход Мауэ, рассмотрел один из типов электронных поверхностных состояний, а именно состояния типа Шокли, на основе метода псевдопотенциала.

В начале 60-х годов эксперименты по дифракции электронов низких энергий (ДЭНЭ)¹⁾ показали, что некоторые из свободных поверхностей испытывают перестройку (например, поверхность {111} у Si, сколотая в условиях сверхвысокого вакуума). Эта перестройка (она наблюдается как дополнительные рефлексы на дифракционной картине) оказывается нестабильной и зависит от температуры. Все говорило о том, что поверхность кремния испытывает *сильное* возмущение, которым отнюдь нельзя *пренебречь*, как предполагалось Шокли и другими. Напротив, как показали такие же эксперименты, ионные (т. е. электрически нейтральные) поверхности соединений элементов III—V, II—VI и I—VII групп не перестраиваются (§ 15). Примерами таких поверхностей служат поверхность (100) у NaCl, (110) у цинковой обманки и (10 $\bar{1}$ 0) у вюрцита. Теоретические работы, посвященные изучению поверхностных состояний в кристаллах соединений этого типа и, в частности, в ионных кристаллах, появились только в течение нескольких последних лет. Методы, которые при этом использовались (они будут обсуждаться в дальнейшем), включают методы ненасыщенных связей [28, 29], константы Маделунга [30], ЛКАО [31—33] и потенциала Матье [34].

¹⁾ В отечественной литературе встречается также термин «дифракция медленных электронов» (ДМЭ). — *Прим. перев.*

Глава 3

МЕТОДЫ КРИСТАЛЛИЧЕСКИХ ОРБИТАЛЕЙ

Одним из самых ранних и вместе с тем наиболее плодотворных подходов в теории поверхностных состояний является метод кристаллических орбиталей. Заменяя кристаллический потенциал системой кулоновских и резонансных интегралов, метод кристаллических орбиталей позволяет свести одноэлектронное дифференциальное уравнение Шредингера к эквивалентному разностному уравнению. Приповерхностная область отличается в этом случае от объема изменением величины соответствующих кулоновских интегралов и отсутствием некоторых связей (и отвечающих им резонансных интегралов). Будучи весьма универсальным, метод кристаллических орбиталей может быть применен к широкому кругу вопросов, касающихся поверхностных состояний. Кроме электронного спектра, метод разностных уравнений позволяет исследовать, например, колебательные или фоновые спектры твердых тел [35].

Прежде всего мы опишем прямой способ решения разностных уравнений на языке блоховских функций [36—39]. В общем случае трех измерений явное решение не может быть получено [1], и это заставляет обратиться к более мощному методу резольвенты. Поскольку этот последний недавно подробно обсуждался в работе Коутецкого [2], мы здесь рассмотрим только один иллюстрирующий его пример. Это будет сделано после того, как мы завершим формулировку основных положений теории.

§ 1. Прямой подход

1. Общая теория

Записанное в дираковских обозначениях [40], одноэлектронное уравнение Шредингера имеет вид

$$(H - E) |\psi\rangle = 0, \quad (1.1)$$

где H — эффективный гамильтониан, $|\psi\rangle$ — волновая функция, отвечающая энергии E . Представим функцию $|\psi\rangle$ в виде линейной комбинации атомных орбиталей (ЛКАО) $|\phi_n\rangle$, т. е. положим

$$|\psi\rangle = \sum_n c_n |\phi_n\rangle. \quad (1.2)$$

Здесь функция $|\phi_n\rangle$ описывает движение электрона в поле только атома с номером n , и c_n — коэффициенты ЛКАО. Подставляя выражение (1.2) в уравнение (1.1) и умножая его слева на $\langle\phi_m|$, получаем систему линейных однородных уравнений для коэффициентов c_n :

$$\sum_n (H_{mn} - ES_{mn}) c_n = 0, \quad (1.3)$$

причем

$$H_{mn} = \langle\phi_m|H|\phi_n\rangle, \quad (1.4)$$

$$S_{mn} = \langle\phi_m|\phi_n\rangle. \quad (1.5)$$

Прежде чем решать систему (1.3) для кристалла с одним атомом на ячейку, сделаем два упрощающих предположения. Предположим, во-первых, что атомные орбитали $|\phi_n\rangle$ образуют ортогональную систему так, что матрица перекрытия имеет вид

$$S_{mn} = \delta_{mn} = \begin{cases} 1, & m = n, \\ 0, & m \neq n. \end{cases} \quad (1.6)$$

Во-вторых, примем, что при $m = n$ матричные элементы гамильтониана постоянны, а при $m \neq n$ отличны от нуля и постоянны только матричные элементы, отвечающие взаимодействию между ближайшими соседями. Математически эти предположения, составляющие основу хорошо известного *приближения сильной связи* [37—39] или *приближения Хюккеля* [41, 42], можно записать как

$$H_{mn} = \begin{cases} \alpha, & m = n, \\ \beta, & m = n \pm 1, \\ 0, & \text{в остальных случаях.} \end{cases} \quad (1.7)$$

Здесь α и β обозначают соответственно *кулоновский* и *резонансный интегралы*. С помощью формул (1.6) и

(1.7) уравнение (1.3) приобретает вид

$$2Xc_n = c_{n+1} + c_{n-1}, \quad (1.8)$$

$$2X = \frac{E - \alpha}{\beta}, \quad (1.9)$$

причем величина X имеет смысл *приведенной* энергии. Поскольку α выражает эффективную энергию электрона на атомной орбитали, а β есть энергия связи между ближайшими соседними атомами, то величина X в уравнении (1.9) *безразмерна*.

Итак, мы преобразовали одноэлектронное уравнение Шредингера (1.1) в однородное линейное разностное уравнение второго порядка (1.8) с постоянными коэффициентами [43, 44], пренебрегая перекрытием и принимая приближение сильной связи в методе ЛКАО. Оставшаяся часть этого параграфа будет посвящена нахождению решений уравнения типа (1.8), удовлетворяющих тем или иным граничным условиям. Эти граничные условия выражают специфику данной конкретной системы. Заметим еще, что для нормированных решений

$$\langle \psi | \psi \rangle = 1, \quad (1.10)$$

и коэффициенты c_n , согласно (1.2) и (1.6), должны удовлетворять условию

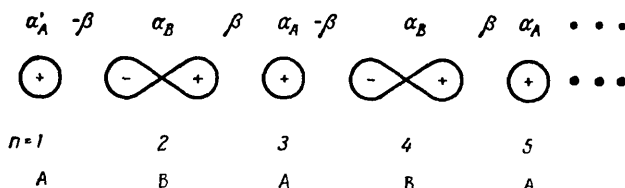
$$\sum_n |c_n|^2 = 1. \quad (1.11)$$

2. Объемные и поверхностные состояния в кристаллах типа АВ

а. Постановка задачи

Выбор в качестве модели кристалла типа АВ позволяет описывать двухатомный или ионный кристалл. Кроме того, имеется возможность свести эти результаты к полученным в случае одноатомной системы. Простейший способ построения кристалла типа АВ состоит в том, что берется линейная цепочка равноотстоящих узлов решетки, пронумерованных, скажем, от 1 до N , и в каждом четном или нечетном узле помещается атом А или В соответственно. Если некоторая $s(p)$ -орбиталь

принадлежит каждому нечетному (четному) атому, то атомам А и В можно приписать кулоновские интегралы α_A, α_B , а связи между соседними атомами — резонансный интеграл β , знак которого чередуется вдоль цепочки



Фиг. 1. Модель орбиталей полубесконечного линейного ионного кристалла.

(фиг. 1). В этом случае уравнение (1.8) можно заменить уравнениями [33]

$$\begin{aligned} (X - z_1) c_n &= c_{n-1} - c_{n+1}, & n \text{ нечетное,} \\ (X + z_1) c_n &= c_{n+1} - c_{n-1}, & n \text{ четное,} \end{aligned} \quad (1.12)$$

где

$$X = \frac{E - \bar{\alpha}}{\beta}, \quad \bar{\alpha} = \frac{\alpha_A + \alpha_B}{2} \quad (1.13)$$

и

$$z_1 = \frac{\alpha_A - \alpha_B}{2\beta} \quad (1.14)$$

— параметр состава ¹⁾.

В полубесконечном кристалле ($N \rightarrow \infty$) уравнения (1.12) справедливы при $n > 1$ (2), если кристалл оканчивается атомом А (В)-типа. Предполагается, что наличие поверхности при $n = 1$ (2) возмущает электронное окружение одного *только* крайнего атома. Поверхностный эффект может быть поэтому учтен простой заменой кулоновского интеграла $\alpha_A(\alpha_B)$ атома с номером $n = 1$ (2) величиной $\alpha'_A(\alpha'_B)$. Граничные условия, которые

¹⁾ Удобно считать ион А (В) электроположительным (электроотрицательным), тогда $\alpha_A > \alpha_B$ и $z_1 > 0$.

получаются тогда из (1.12), принимают вид

$$(X - z_A) c_1 = -c_2, \quad (1.15)$$

$$(X + z_B) c_2 = c_3, \quad (1.16)$$

$$z_A = \frac{\alpha'_A - \bar{\alpha}}{\beta}, \quad z_B = \frac{\bar{\alpha} - \alpha'_B}{\beta}, \quad (1.17)$$

где z_A и z_B — *параметры возмущения*. Далее, в силу полубесконечности кристалла поверхностным эффектом при $n = N$ можно пренебречь [1]:

$$c_m = 0, \quad m = N + 1. \quad (1.18)$$

б. Решения

Решения уравнения (1.12) можно найти, если положить [45]

$$\begin{aligned} c_n &= u^n, & n \text{ четное,} \\ c_n &= v^n, & n \text{ нечетное.} \end{aligned} \quad (1.19)$$

Подставляя (1.19) в (1.12) и исключая v , приходим к уравнению четвертой степени

$$u^4 + \gamma u^2 + 1 = 0, \quad (1.20)$$

где

$$\gamma = X^2 - z_1^2 - 2. \quad (1.21)$$

Выбор

$$\gamma = -2 \cos \theta \quad (1.22)$$

означает, что общее решение (1.12) можно записать в виде

$$c_n = P \cos \frac{n\theta}{2} + Q \sin \frac{n\theta}{2}, \quad n \text{ нечетное,} \quad (1.23)$$

что, согласно (1.18), преобразуется в

$$c_n = R \sin \frac{(m-n)\theta}{2}, \quad n \text{ нечетное,} \quad (1.24)$$

где

$$R = \frac{P}{\sin(m\theta/2)} \quad (1.25)$$

— нормировочный множитель. Подставляя решение (1.24) во второе из уравнений (1.12), получаем

$$c_n = KR \cos \frac{(m-n)\theta}{2}, \quad n \text{ четное}, \quad (1.26)$$

где

$$K = - \frac{2 \sin(\theta/2)}{X + z_1}. \quad (1.27)$$

С помощью выражений (1.21) и (1.22) получаем формулу для соответствующих приведенных энергий [33]

$$X = \pm \left(z_1^2 + 4 \sin^2 \frac{\theta}{2} \right)^{1/2}. \quad (1.28)$$

Неизвестный параметр θ , введенный формулой (1.22), определяется из уравнений для собственных значений, которые вытекают из граничных условий (1.15) и (1.16). Для поверхностей с атомами А и В соответственно имеем

$$(X - z_A) \sin N \frac{\theta}{2} = -K \cos(N-1) \frac{\theta}{2}, \quad (1.29)$$

$$(X + z_B) K \cos(N-1) \frac{\theta}{2} = \sin(N-2) \frac{\theta}{2}. \quad (1.30)$$

Анализ этих уравнений показывает, что все их решения ¹⁾ действительны, за исключением двух, которые в зависимости от значений величин z_1 , А, В могут быть либо действительными, либо комплексными.

в. Объемные состояния

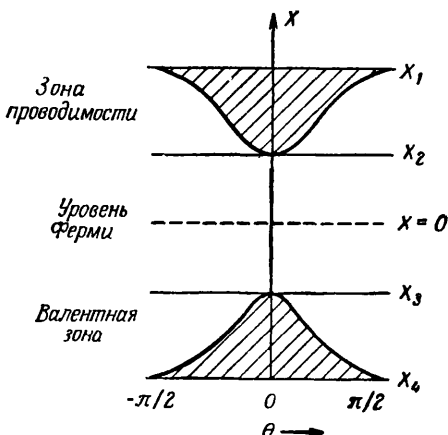
В случае действительного θ коэффициенты (1.24) и (1.26) периодически изменяются внутри кристалла (т. е. состояния нелокализованы), причем амплитуды, отвечающие узлам с атомами А и В, различны. Следовательно, вероятности нахождения электрона в атоме А и в атоме В отличаются. Этого следовало ожидать, поскольку значения кристаллического потенциала в узлах типа А и В неодинаковы, что мы и отметили введением разных коэффициентов α_A и α_B .

¹⁾ Следует рассматривать эти решения только в интервале $0 \leq \theta \leq \pi$, так как в остальной области они просто повторяются.

Приведенные энергии (1.28) располагаются тогда в двух зонах (фиг. 2), ограниченных значениями.

$$X_{1,4} = \pm (z_1^2 + 4)^{1/2}, \quad X_{2,3} = \pm z_1. \quad (1.31)$$

Ширины разрешенных и запрещенной зон легко найти из выражений (1.31). Они оказываются зависящими только от параметра состава z_1 . Две разрешенные зоны располагаются симметрично относительно линии $X = 0$ (уровень Ферми) и разделены запрещенной энергетической



Фиг. 2. Диаграмма приведенной энергетической зонной структуры.

зоной шириной $2z_1$ [46]. В случае кристалла, состоящего из атомов одного сорта $\alpha_A = \alpha_B$ (т. е. $z_1 = 0$), и запрещенная зона исчезает, так что спектр состоит только из одной разрешенной зоны [1, 3, 14, 16, 23].

г. Поверхностные состояния

Для рассмотрения комплексных решений уравнений (1.29) и (1.30) положим [31]

$$\theta = (\zeta + i\mu), \quad \zeta, \mu \text{ (действительны)} > 0. \quad (1.32)$$

Мы выбрали положительное значение μ , чтобы обеспечить экспоненциальное затухание волновых функций

внутри кристалла. Так как энергия X в (1.28) должна быть действительной величиной, то выполняется равенство

$$\operatorname{Im} \sin^2 \frac{\theta}{2} = 0,$$

или

$$\sin \xi \operatorname{sh} \mu = 0. \quad (1.33)$$

Однако $\operatorname{sh} \mu \neq 0$ для $\mu > 0$, и поэтому

$$\xi = k\pi, \quad k = 0, 1, \dots \quad (1.34)$$

Необходимо рассмотреть только решения с $k = 0$ и 1 , поскольку остальные значения k приводят лишь к простому повторению решений. Таким образом, θ принимает вид

$$\theta = i\mu \quad \text{и} \quad \pi + i\mu, \quad (1.35)$$

причем энергии (1.28) теперь равны

$$X = \pm \left(z_1^2 - 4 \operatorname{sh}^2 \frac{\mu}{2} \right)^{1/2}, \quad (1.36)$$

$$X = \pm \left(z_1^2 + 4 \operatorname{ch}^2 \frac{\mu}{2} \right)^{1/2}. \quad (1.37)$$

Уровни (1.36) и (1.37), расположенные соответственно *внутри* запрещенной зоны и *вне* разрешенных зон (фиг. 2), называются *внутренними* и *внешними* уровнями [31, 33]. Кроме того, при $X > 0$ ($X < 0$) говорят об уровнях P (N)-типа, которые не следует путать с примесными уровнями в полупроводниках p - и n -типа.

1) *Внутренние состояния*. При $\theta = i\mu$ и $N \rightarrow \infty$ выражения (1.24) и (1.26) сводятся к виду

$$\begin{aligned} c_n &= PL^{n/2}, & n \text{ нечетное,} \\ c_n &= KPL^{n/2}, & n \text{ четное,} \end{aligned} \quad (1.38)$$

где

$$L = \exp(-\mu), \quad 0 < L < 1 \quad (1.39)$$

имеет тот же смысл, что и параметр λ в работе Дэвисона и Коутецкого [31]. Отрицательный показатель экспоненты в (1.38) означает, что коэффициенты имеют наибольшее значение при $n = 1$ и постепенно убывают с ростом n . Итак, мы можем сказать, что электрон

локализован у поверхности кристалла, причем степень локализации тем выше, чем больше μ .

Если $\theta = i\mu$, то из уравнений (1.29) и (1.30) при помощи соотношений (1.21), (1.22) и (1.27) получаем

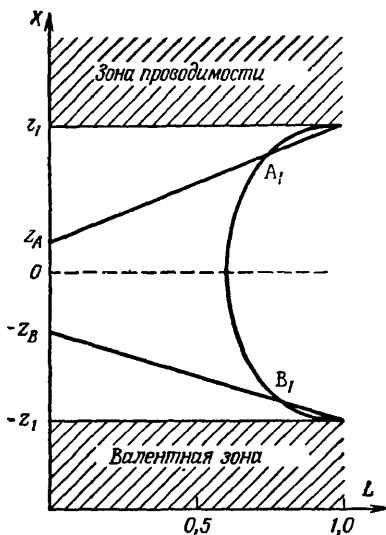
$$X = z_A + (z_1 - z_A)L, \quad (1.40)$$

$$X = -[z_B + (z_1 - z_B)L]. \quad (1.41)$$

Выражение для энергии (1.36), записанное с помощью параметра L , имеет вид

$$X = \pm [z_1^2 - (1 - L)^2 L^{-1}]^{1/2}, \quad (1.42)$$

откуда следует, что положение внутренних уровней на плоскости XL (фиг. 3) отвечает точкам пересечения A_1



Фиг. 3. Положение энергетических уровней внутренних поверхностных состояний А- и В-типа.

и B_1 прямых (1.40) и (1.41) с кривой (1.42) в интервале $0 < L < 1$. Точки пересечения при $X = \pm z_1$, для которых $L = 1$ или $\mu = 0$, — это тривиальные решения, отвечающие краям валентной зоны и зоны проводимости.

Исключая X из (1.42) и (1.40) или (1.41), имеем [31, 33]

$$f_r L^2 + g_r L - 1 = 0, \quad r = A \text{ или } B, \quad (1.43)$$

где

$$f_r = (z_1 - z_r)^2, \quad (1.44)$$

$$g_r = 1 + z_1^2 - z_r^2. \quad (1.45)$$

Корень уравнения (1.43) при условии $L > 0$ имеет вид

$$L = \frac{-g_r + (g_r^2 + 4f_r)^{1/2}}{2f_r}. \quad (1.46)$$

Подставляя это выражение в (1.40), (1.41) или (1.42), мы можем получить энергию X как явную функцию параметров состава и параметров возмущения. Согласно определению (1.17), формуле (1.46) и из фиг. 3 видно, что при $L > 0$ должно быть

$$z_1 > z_r, \quad r = A \text{ или } B. \quad (1.47)$$

Эти два последних неравенства выражают *условия существования* или *пороговые условия* для возникновения внутренних состояний. Другими словами, при их выполнении обеспечено пересечение кривых в точках A_1 и B_1 на фиг. 3. В одноатомном кристалле $z_1 = 0$, и запрещенной зоны внутри основной разрешенной зоны нет. Тогда неравенство (1.47) не выполнено и внутренние состояния исчезают, сливаясь с объемными.

Если выразить величину L через значения энергии, отвечающие краям зоны проводимости и валентной зоны, то, согласно (1.40) и (1.41), имеем

$$L = \frac{(z_1 - z_r) - (z_1 - X)}{z_1 - z_r}. \quad (1.48)$$

Физически можно считать, что поверхностные состояния типа A и B обладают энергией «захвата» $z_1 - X$ и «потенциальной энергией» $z_1 - z_r$. Числитель в выражении (1.48) дает тогда «нулевую энергию» электрона, захваченного в поверхностную ловушку, при этом

$$L = \frac{\text{Нулевая энергия}}{\text{Потенциальная энергия}}.$$

Поскольку $0 < L < 1$, нулевая энергия меньше, чем потенциальная. Для очень мелких поверхностных состояний величина f_r , согласно (1.44), мала, и выражение с квадратным корнем в (1.46) можно разложить в ряд, что дает

$$L \sim g_r^{-1}. \quad (1.49)$$

Отсюда, если учесть (1.45), видно, что отношение нулевой энергии к потенциальной будет наименьшим при $z_1 \gg z_r$ ¹⁾. Именно такая ситуация обнаружена в кристаллах с широкой запрещенной зоной (шириной $2z_1$) и глубокими поверхностными состояниями, потенциальная энергия которых порядка $z_1 - z_r$.

2) *Внешние состояния.* Пусть теперь $\theta = (\pi + i\mu)$. Соответствующие уравнения можно получить из уравнений для внутренних состояний простой заменой L на $-L$ [31]. Таким образом, вместо (1.38) имеем

$$\begin{aligned} c_n &= P(-1)^{n/2} L^{n/2}, & n \text{ нечетное,} \\ c_n &= KP(-1)^{n/2} L^{n/2}, & n \text{ четное,} \end{aligned} \quad (1.50)$$

что совпадает, за исключением множителя $(-1)^{n/2}$, с выражением для экспоненциально затухающих коэффициентов внутренних состояний. Уравнения для собственных значений (1.40) и (1.41) теперь имеют вид

$$X = z_A - (z_1 - z_A)L, \quad (1.51)$$

$$X = -[z_B - (z_1 - z_B)L], \quad (1.52)$$

а вместо выражения для приведенной энергии (1.42) имеем

$$X = \pm [z_1^2 + (1 + L)^2 L^{-1}]^{1/2}. \quad (1.53)$$

Графически положение внешних уровней можно найти, если построить график X как функции L , согласно уравнениям (1.51) и (1.52) вместе с кривой (1.53) (фиг. 4). Искомые значения L (и, следовательно, μ) от-

¹⁾ При условии $z_1 \gg z_r$ величину f_r нельзя, вообще говоря, считать малой, однако разложение все же законно, поскольку в этом случае величины g_r и f_r одного порядка и обе много больше 1, так что $g_r^2 \gg 4f_r$. — *Прим. перев.*

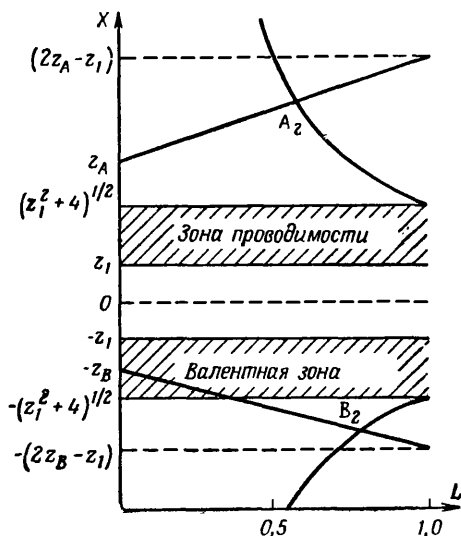
вечают точкам пересечения A_2 и B_2 этих кривых или положительным корням уравнений [ср. (1.43)]

$$f_r L^2 - g_r L - 1 = 0, \quad \text{где } r = A \text{ или } B, \quad (1.54)$$

т. е. [ср. (1.46)]

$$L = \frac{g_r + (g_r^2 + 4f_r)^{1/2}}{2f_r}. \quad (1.55)$$

Подставляя теперь (1.55) в формулы (1.51), (1.52) или (1.53), получаем X как функцию параметров z .



Фиг. 4. Положение энергетических уровней внешних поверхностных состояний А- и В-типа.

Так как $L < 1$, формула (1.55) дает условие существования [31]

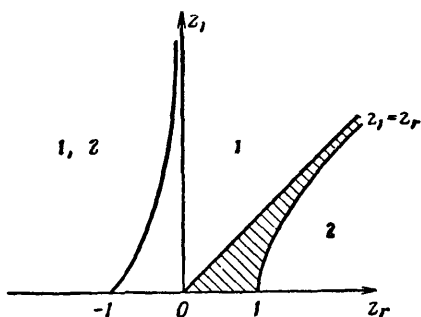
$$z_r(z_r - z_1) > 1, \quad r = A \text{ или } B. \quad (1.56)$$

Это же условие можно также получить, если заметить, что при $L = 1$ энергия поверхностного состояния, равная, согласно (1.51) или (1.52), $|2z_r - z_1|$, должна лежать вне зоны объемных состояний, край которой

отвечает энергии $|z_1^2 + 4|^{1/2}$ [см. (1.53)]. Положив $z_1 = 0$, мы переходим от (1.56) к хорошо известному условию существования поверхностных состояний в одномерном, одноатомном кристалле [1, 3, 16, 23]:

$$z_r^2 > 1. \quad (1.57)$$

Если теперь $z_1 > 0$, то при положительных z_r наличие члена с z_1 в неравенстве (1.56) означает, что внешние поверхностные состояния с большей вероятностью возникают в одноатомных кристаллах, чем в кристаллах



Фиг. 5. Схема областей существования внутренних (1) и внешних (2) состояний на плоскости $z_1 z_r$. Объемным состояниям соответствует заштрихованный участок.

типа АВ. Это же справедливо и в случаях больших z_1 , т. е. для веществ с широкой запрещенной зоной, атомы компонент которых сильно различаются. Напротив, при $z_r < 0$ внешние состояния чаще встречаются в кристаллах типа АВ, чем в одноатомных. В этом случае при больших z_1 поверхностные состояния возникают в кристаллах типа АВ даже при $z_r^2 < 1$.

На фиг. 5 выделены области плоскости $z_1 z_r$, в которых выполнены или нарушены неравенства (1.47) и (1.56) [31]. При значениях параметров (z_1, z_r) , отвечающих заштрихованному участку, существуют *только* объемные состояния. При этом нарушаются *оба* неравенства (1.47) и (1.56). Распределение внутренних и внешних состояний в трех других областях плоскости обозначено соответственно цифрами 1 и 2.

На опыте легче наблюдать внутренние состояния, чем внешние. Это происходит потому, что уровни, лежащие в запрещенной зоне между зоной проводимости и валентной зоной полупроводника, могут заметно изменить положение уровня Ферми ($X = 0$). С точки зрения теории, однако, полезно рассмотреть состояния обоих типов.

Поверхностные состояния в сплавах типа АВ изучались с помощью метода матрицы рассеяния в работе [47]. Этот метод будет обсуждаться в п. 1 § 3.

3. Другие работы по бинарным кристаллам

Выше мы умышленно сосредоточили внимание на детальном анализе объемных и поверхностных состояний в бинарных (в частности, ионных) кристаллах типа АВ. Это было сделано с той целью, чтобы читатель составил четкое представление о *применении* прямого метода кристаллических орбиталей к решению некоторой частной проблемы. В настоящем разделе мы кратко рассмотрим ряд других задач, которые относятся к бинарным кристаллам и были исследованы с помощью указанного метода.

а. Хемосорбционные состояния

Наряду с рассмотрением объемных и поверхностных состояний Дэвисон и Коутецкий [31] обсудили также вопрос о хемосорбционных состояниях в кристалле типа АВ. Последние возникают, когда чужеродный атом (адсорбат) взаимодействует с поверхностью кристалла посредством образования химической связи между атомом адсорбата и кристаллом. Для кристалла, ограниченного атомом А, граничные условия наряду с (1.18) будут

$$(X - z_1 - z_2)c_1 = c_2 + \eta c_v, \quad (1.58)$$

$$(X - z_1 - z^*)c_v = c_1, \quad (1.59)$$

где c_v — коэффициент, отвечающий адсорбированному атому v , и

$$z_2 = \frac{\alpha'_A - \alpha_A}{\beta}, \quad z^* = \frac{\alpha^* - \alpha_A}{\beta}, \quad \eta = \frac{\beta^*}{\beta}, \quad (1.60)$$

причем α^* — кулоновский интеграл чужого атома, а β^* — резонансный интеграл его связи с кристаллом. Кроме коэффициентов, соответствующих объемным состояниям, новое граничное условие (1.59) дает выражение для коэффициента c_v , отличающееся от формул для коэффициентов в объеме. Хотя в основном спектр приведенной энергии остается тем же [т. е. краям зон отвечает выражение (1.31) и могут возникать внутренние и внешние состояния], граничное условие (1.58) дает новое уравнение для собственного значения θ . Поэтому каждому отдельному объемному или хемосорбционному состояниям отвечают разные уровни в энергетическом спектре. Условия существования последних имеют теперь более сложный вид, чем в случае чистой поверхности. В пределе, когда z_1 стремится к нулю, граничные условия (1.58) и (1.59) описывают процесс адсорбции на поверхности атомного кристалла [1, 2, 8, 23, 48, 49].

6. Деформированные кристаллы

Как отмечали Раймер [50] и Жданов [51], существуют теоретические и экспериментальные доказательства того, что расположение атомов у поверхности отличается от их расположения в объеме. Точные эксперименты по ДЭНЭ на поверхностях германия [52] и алмаза [53] подтверждают эту точку зрения. Влияние поверхностного возмущения решетки на поверхностные состояния в кристаллах типа АВ изучалось теоретически в рамках прямого метода кристаллических орбиталей [54]. Однако в случае одноатомных кристаллов применялись как метод кристаллических орбиталей [25, 55], так и метод кристаллического потенциала [56—58]. Современное развитие проблемы таммовских состояний у перестроенных кристаллических поверхностей было недавно рассмотрено Дэвисоном и Стенслицкой [59].

Граничные условия для деформированного кристалла АВ-типа можно легко получить из (1.58) и (1.59), полагая $z_2 = 0$ и переопределяя некоторые параметры [54]. При этом c_v — коэффициент разложения волновой функции, отвечающий первому атому в кристалле, связанному с остальным кристаллом связью (β^*). Длина этой

связи отличается от длины связей (β) между атомами в недеформированной части кристалла. Деформация, которую претерпевает кристалл при образовании поверхности, проявляется в изменении функционального вида коэффициента c_v по сравнению с коэффициентами, отвечающими атомам в объеме. Пожалуй, одним из наиболее интересных аспектов работы, посвященной деформированным кристаллам типа АВ, являются результаты, полученные при раздельном учете поверхностной деформации (η) и поверхностного возмущения (z^*). Среди этих результатов укажем следующие. 1) Положения краев зон (1.31) не зависят от замены поверхностного возмущения деформацией поверхности, так как определяются только составом кристалла. 2) В присутствии только поверхностной деформации внутренних состояний не существует (т. е. возможны только внешние состояния). 3) Внутренние состояния вызываются только поверхностным возмущением, тогда как внешние могут быть обусловлены как поверхностным возмущением, так и деформацией. 4) При наличии и поверхностного возмущения, и поверхностной деформации положение энергетических уровней внутренних состояний нечувствительно к влиянию деформации, но определяется исключительно влиянием поверхностного возмущения. Напротив, положение внешних уровней зависит и от возмущения, и от деформации. Заметим, что здесь рассматриваются только смещения, нормальные к поверхности.

Коутецкий [25], пользуясь резольвентным методом кристаллических орбиталей (§ 2), изучал поверхностные состояния одноатомного кристалла, в котором потенциал поверхностного возмущения действует в глубь кристалла вплоть до $(Z + 1)$ -й элементарной ячейки. Оказалось, что число возникающих поверхностных состояний может достигать максимум до $(Z + 1)$. При $Z = 1$ Коутецкий нашел, что $|c_v/c_1| < 1$, так что наибольший коэффициент отвечает не первому поверхностному атому, а *следующему* за ним. Энергия таких *подповерхностных* состояний может оказаться ниже, чем энергия обычных поверхностных. Фаризо [58] подтвердил эти результаты, пользуясь методом функций Грина (п. 2 § 3) и рассматривая возмущенную модель Кронига — Пенни [13].

Поверхностные состояния в одноатомных кристаллах были также интенсивно исследованы Стенсличкой, применявшей как прямой метод кристаллических орбиталей [55], так и метод матрицы рассеяния [57, 58]. Особо здесь следует отметить расчеты по прямому методу кристаллических орбиталей для трехмерной модели Li, показавшие, что если поверхностная деформация уменьшает постоянную решетки у поверхности на величину $\epsilon \geq 0,7 \text{ \AA}$ (или с учетом перекрытия $\epsilon \geq 0,30 \text{ \AA}$), то в энергетическом спектре кристалла появляется дополнительная зона поверхностных состояний. Кроме того, поскольку поверхностная деформация заметно влияет на существование поверхностных состояний, то работа Стенсличкой дает новое подтверждение предположению Хенслера [60] о происхождении поверхностных состояний. Последний отмечал, что особое расположение поверхностных атомов по отношению к их ближайшим соседям может быть в основном ответственно за появление наблюдаемых состояний. Экспериментальной стороне вопроса о поверхностной деформации будет посвящено более подробное обсуждение в гл. 7.

в. Кристаллы типа A_rB_s

Объемные состояния ряда бинарных систем были детально изучены Гоффманом [46], который применил технику производящей функции в рамках прямого метода кристаллических орбиталей при решении системы связанных линейных разностных уравнений, описывающих электронные состояния рассматриваемой системы. Подобно случаю кристалла типа АВ число уравнений в системе равно числу атомов $(r + s)$ в элементарной ячейке кристалла, которое в свою очередь равно числу зон энергетического спектра. Ширины зон объемных состояний оказались зависящими от состава кристалла. Кроме того, Гоффман вывел формулы для плотности состояний, изучил влияние интеграла перекрытия и обобщил [61] результаты для кристаллов типа АВ на случай двух и трех измерений.

Полный энергетический спектр (объемные и поверхностные уровни) бинарных сплавов, рассмотренных Гоффманом, был исследован Дэвисоном, Лингом и Гошем [62] с помощью прямого метода кристаллических орбиталей. Было показано, что вне $(r + s)$ зон могут возникнуть самое большее $2(r + s)$ поверхностных уровней. Полученные уравнения определяют положение и число этих уровней. Кроме того, если число $(r + s)$ нечетное (четное), то, как было показано, поверхностные состояния могут (не могут) возникать во всех (в некоторых) запрещенных зонах. При этом отмечалось, что, поскольку в работе рассматривались только кристаллы, оканчивающиеся атомом типа А, были найдены только $(r + s)$ поверхностных уровней, остальные $(r + s)$ уровней возникают в кристалле, оканчивающемся атомом типа В (п. 2 § 1).

Указанный анализ A_rB_s -сплавов привел к построению теории электронных состояний в кристаллах соединений общего вида, основанной на резольвентном методе кристаллических орбиталей [32]. Элементарная ячейка таких кристаллов содержит произвольное число различных атомов, расположенных в каком угодно порядке и связанных друг с другом силами различной интенсивности. Такая постановка задачи является достаточно общей, чтобы охватить все предыдущие случаи.

§ 2. Метод резольвенты

1. Общая теория

Рассмотрим систему (молекулярную или кристаллическую) с эффективными одноэлектронными гамильтонианами: H — с учетом взаимодействия между частями системы, H^0 — без учета этого взаимодействия, причем

$$H = H^0 + V, \quad (2.1)$$

где V — потенциал взаимодействия. Если собственный вектор и собственное значение оператора H^0 для j -го состояния соответственно $|\psi_j^0\rangle$ и E_j^0 , то уравнение (1.3) для невзаимодействующей (невозмущенной) системы

можно записать в виде

$$\sum_l H_{kl}^0 c_{lj}^0 = E_j^0 \sum_l \delta_{kl} c_{lj}^0, \quad (2.2)$$

где c_{lj}^0 — коэффициенты разложения

$$|\psi_j^0\rangle = \sum_l c_{lj}^0 |\phi_l^0\rangle, \quad (2.3)$$

аналогичного разложению (1.2). Если считать, что векторы $|\psi_j^0\rangle$ и $|\phi_l^0\rangle$ ортогональны, то должно выполняться условие

$$\sum_m c_{km}^{0+} c_{ml}^0 = \delta_{kl}, \quad (2.4)$$

или, другими словами, коэффициенты c_{km}^0 служат элементами *унитарной* матрицы [63]. Пользуясь соотношением (2.4), перепишем уравнение (2.2) в виде

$$\sum_l H_{kl}^0 c_{lj}^0 = E_j^0 \sum_m c_{km}^0 \delta_{mj}$$

или, в матричной записи,

$$\mathbf{H}^0 \mathbf{c}^0 = \mathbf{c}^0 \mathbf{E}^0, \quad (2.5)$$

где $\mathbf{c}^0$ — матрица, столбцы которой образуют собственные векторы оператора $\mathbf{H}^0$, а $\mathbf{E}^0$ — диагональная матрица, элементы которой E_j^0 дают соответствующие собственные значения.

Для взаимодействующей (возмущенной) системы уравнение (2.5) перепишется в следующем виде¹⁾:

$$(\mathbf{H}^0 + \mathbf{V}) \mathbf{c} = \mathbf{E} \mathbf{1} \mathbf{c} \quad (2.6)$$

или, после некоторых преобразований,

$$\mathbf{c} = (\mathbf{E} \mathbf{1} - \mathbf{H}^0)^{-1} \mathbf{V} \mathbf{c}. \quad (2.7)$$

Можно показать, пользуясь соотношениями (2.4) и (2.5), что

$$(\mathbf{E} \mathbf{1} - \mathbf{H}^0)^{-1} = \mathbf{c}^0 (\mathbf{E} \mathbf{1} - \mathbf{E}^0)^{-1} \mathbf{c}^{0+}. \quad (2.8)$$

¹⁾ Здесь $\mathbf{1}$ — символ единичной матрицы (оператора). — *Прим. перев.*

Подставляя это выражение в (2.7), имеем [64]

$$c = G V c, \quad (2.9)$$

где матрицу

$$G = c^0 (E I - E^0)^{-1} c^{0+} \quad (2.10)$$

называют *резольвентной* (или *гриновской*) *матрицей*. Если считать, что в диагональной матрице V отличны от нуля элементы $V_1, \dots, V_M$, то тогда условием существования нетривиальных решений уравнения (2.9) будет

$$\begin{vmatrix} V_1 G_{11} - 1 & V_2 G_{12} & \dots & V_M G_{1M} \\ V_1 G_{21} & V_2 G_{22} - 1 & \dots & V_M G_{2M} \\ \vdots & & & \\ V_1 G_{M1} & V_2 G_{M2} & \dots & V_M G_{MM} - 1 \end{vmatrix} = 0. \quad (2.11)$$

Решение этого уравнения с учетом (2.10) позволяет выразить возмущенные функции через известные невозмущенные.

Описанный метод, развитый первоначально рядом авторов [65—67], был затем обобщен Коутецким [2], который применил формализм операторов проектирования [68]. Оператору проектирования T ставится здесь в соответствие пространство функций t , на которые он проектирует. Так как оператор T всей системы составлен из операторов T_i , отвечающих ее подсистемам, то

$$T = \cup T_i, \quad (2.12)$$

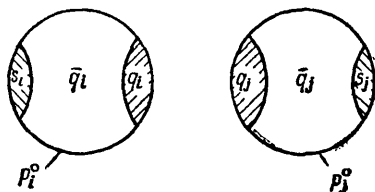
где $\cup$ означает операцию «объединения». Графически не взаимодействующую и взаимодействующую системы можно описать диаграммами, представленными на фиг. 6. Для невзаимодействующей системы собственные функции оператора Гамильтона H^0 образуют базис функционального пространства p^0 , подпространства которого обозначим как p_i^0 . Пользуясь языком операторов, положим

$$p^0 = \cup p_i^0 = I, \quad (2.13)$$

где I — единичный оператор. Таким образом, оператор P^0 проектирует на полную систему функций. Далее, для $i \neq j$

$$P_i^0 P_j^0 = 0, \quad (2.14)$$

т. е. операторы проектирования $P_{i,j}^0$ проектируют на неперекрывающиеся подпространства $p_{i,j}^0$, если $i \neq j$.



Фиг. 6. Схематическое изображение i -го и j -го подпространств взаимодействующей и невзаимодействующей систем.

В случае системы с взаимодействием пространство p собственных функций гамильтониана H из (2.1) есть (см. фиг. 6)

$$p = p^0 - s, \quad (2.15)$$

таким образом, пространство p можно считать подпространством, входящим в состав p^0 . Удалить пространство s из p^0 с образованием пространства p — это все равно, что, например, вынуть один атом из циклической цепочки атомов, представляющей бесконечный кристалл. В результате мы получим конечную (или полубесконечную) цепочку. В общем случае пространство p можно разбить на сумму двух пространств

$$p = q + \bar{q}. \quad (2.16)$$

Здесь пространство $\bar{q}$ объединяет все состояния, для которых матричные элементы V равны нулю, а q — пространство, дополнительное к $\bar{q}$. Возвращаясь к нашему примеру, можно сказать, что образование полубесконечного кристалла (p) приводит к наличию поверхности, что возмущает кристаллический потенциал некоторой части или всей области, отвечающей пространству q . При этом остальная часть ($\bar{q}$) кристалла остается совер-

шенно невозмущенной. Отсюда сразу получаем операторное равенство

$$\bar{Q}V = V\bar{Q} = 0, \quad (2.17)$$

и, как следует из (2.13), (2.15) и (2.16),

$$I = S + Q + \bar{Q}. \quad (2.18)$$

Теперь следует решить уравнение для собственных значений

$$H|\psi\rangle = (H^0 + V)|\psi\rangle = E|\psi\rangle. \quad (2.19)$$

Его можно переписать [ср. с (2.7) и (2.9)] в виде

$$|\psi\rangle = GIVI|\psi\rangle, \quad (2.20)$$

где [69] [ср. с (2.8) и (2.10)]

$$G = (EI - H^0)^{-1} = \frac{\sum_i |\psi_i^0\rangle\langle\psi_i^0|}{E - E_i^0}, \quad (2.21)$$

причем $|\psi_i^0\rangle\langle\psi_i^0|$ — оператор проектирования на j -ю молекулярную орбиталь. Подставляя (2.18) в (2.20), с помощью (2.17) получаем [2, 32]

$$|\psi\rangle = G(QV - SH^0)Q|\psi\rangle, \quad (2.22)$$

так как

$$S|\psi\rangle = 0 \quad (2.23)$$

и

$$SH|\psi\rangle = SH^0|\psi\rangle + SV|\psi\rangle = 0. \quad (2.24)$$

Уравнение для амплитуд, аналогичное (2.9), будет иметь вид

$$\langle\phi_i^0|\psi\rangle = \langle\phi_i^0|G(QV - SH^0)Q|\psi\rangle, \quad (2.25)$$

причем

$$\begin{aligned} \langle\phi_i^0|\psi\rangle &\neq 0, & |\phi_i^0\rangle &\in p, \\ \langle\phi_i^0|\psi\rangle &= 0, & |\phi_i^0\rangle &\in s. \end{aligned} \quad (2.26)$$

Преимущество от введения слагаемого с S в уравнении (2.25) состоит в возможности рассмотрения такой исходной системы, которая больше и проще (например, по

своим свойствам симметрии), чем система, фактически подлежащая изучению. Условие существования нетривиальных решений уравнения (2.25) гласит [ср. (2.11)]:

$$\det \{ \langle \phi_l^0 | G(QV - SH^0)Q - I | \phi_k^0 \rangle \} = 0, \quad (2.27)$$

причем порядок детерминантного уравнения определяется числом состояний l , удовлетворяющих $l \in Q$. В заключение отметим, что существует связь [32] между методом резольвенты и методом разделения [11, 70, 71].

2. Квантовые состояния в кристалле в приближении взаимодействия вторых ближайших соседей

а. Решение для циклического кристалла

Рассмотрим трехмерный циклический одноатомный кристалл. Пусть σ , τ , ν — взаимно перпендикулярные векторы элементарных трансляций. Одноэлектронную волновую функцию в таком кристалле можно записать в виде

$$| \psi_{\theta\eta\nu}^0(\mathbf{r}) \rangle = N^{-1/2} \sum_{l=0}^{N-1} e^{il\theta} | a_{\eta\nu}^0(\mathbf{r} - l\sigma) \rangle, \quad (2.28)$$

причем

$$| a_{\eta\nu}^0(\mathbf{r} - l\sigma) \rangle = N^{-1} \sum_{m,n=0}^{N-1} e^{i(m\eta+n\nu)} | a^0(\mathbf{r} - l\sigma - m\tau - n\nu) \rangle \quad (2.29)$$

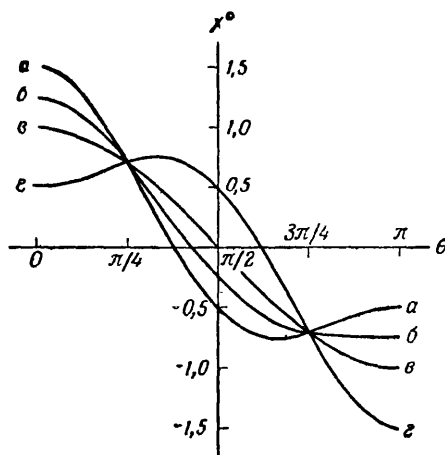
— одномерные функции Ванье [67, 72, 73] или слоистые орбитали [32], а N — число элементарных ячеек в направлении некоторой элементарной трансляции. Циклические граничные условия Борна — Кармана означают, что

$$\theta, \eta, \nu = \frac{2\pi t}{N}, \quad 0 \leq t \leq N-1. \quad (2.30)$$

Сравнение (2.28) и (2.3) показывает, что

$$c_l^0 = e^{il\theta}, \quad (2.31)$$

и слоистые орбитали в трехмерном случае отвечают атомным орбиталям линейного кристалла. Если кроме матричных элементов $H_{n, n\pm 1}$ в определении (1.7) исчезающими оказываются также матричные элементы $H_{n, n\pm 2} = \beta_2$, отвечающие вторым ближайшим соседям,



Фиг. 7. Зависимость приведенной X^0 от θ .

$a - \rho = 0,5$; $b - \rho = 0,25$; $c - \rho = 0$; $d - \rho = -0,5$.

то уравнение (1.8) заменяется разностным уравнением четвертого порядка [2, 4, 67, 74—79]:

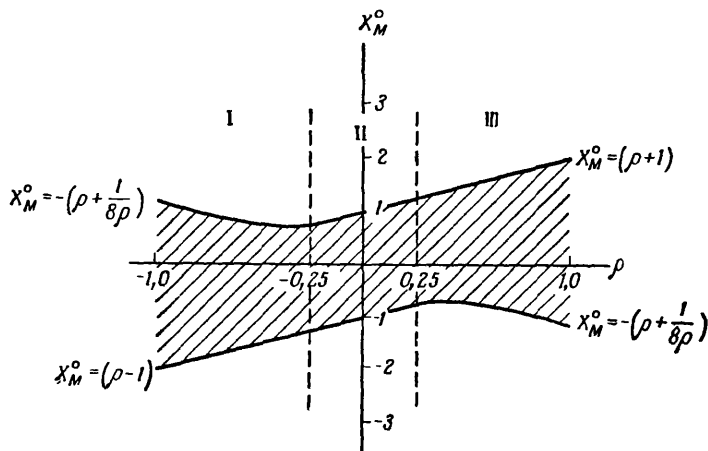
$$2X^0 c_i^0 = (c_{i+1}^0 + c_{i-1}^0) + \rho(c_{i+2}^0 + c_{i-2}^0), \quad (2.32)$$

где $|\rho| = |\beta_2/\beta_1| < 1$ определяет величину взаимодействия вторых ближайших соседей. Из (2.31) и (2.32) получаем, что

$$X^0 = \cos \theta + \rho \cos 2\theta. \quad (2.33)$$

График X^0 как функция θ показан на фиг. 7 для интервала $0 \leq \theta \leq \pi$ при различных значениях параметра ρ . Случай $\rho = 0$ отвечает приближению сильной связи. При $|\rho| > 1/4$ кривые более не являются монотонными, как в случае приближения сильной связи, но обладают относительными экстремумами. Экстремальное значение

$\chi^0(\chi_M^0)$ в функции ρ показано на фиг. 8. Как видно, энергетическую зону можно разделить на три области в соответствии со значением параметра ρ . При $\rho \rightarrow 0$ область II вырождается в обычную зону приближения



Фиг. 8. Зависимость положения краев зоны объемных состояний от параметра взаимодействия вторых ближайших соседей (ρ).

сильной связи. Соображения симметрии для *точного* решения уравнения Шредингера требуют, чтобы относительные экстремумы исчезали.

6. Решения для полубесконечного кристалла

Если пронумеровать N атомов циклического кристалла (ρ^0), используя индекс n ($0 \leq n \leq N-1$), то конечный кристалл (ρ) в приближении взаимодействия вторых ближайших соседей можно получить, удаляя $(N-1)$ -й и $(N-2)$ -й атомы (s). Следовательно,

$$S = \sum_{l=N-2}^{N-1} |a_l^0\rangle \langle a_l^0|. \quad (2.34)$$

Для полубесконечного кристалла $N \rightarrow \infty$, и волновую функцию можно разложить в бесконечный ряд по сло-

истым орбиталям:

$$|\psi_{\eta\nu}(\mathbf{r})\rangle = \sum_{l=0}^{\infty} c_{\eta\nu, l} |a_{\eta\nu}^0(\mathbf{r} - l\sigma)\rangle. \quad (2.35)$$

Образование поверхности при $l = 0^1$) приводит к возникновению потенциала поверхностного возмущения V , проникающего в кристалл только на глубину первой ячейки за плоскость поверхности, вектор нормали к которой σ . Матричный элемент потенциала возмущения может быть поэтому записан в виде

$$\langle a_{\eta\nu}^0(\mathbf{r} - l\sigma) | V | a_{\eta'\nu'}^0(\mathbf{r} - l'\sigma) \rangle = A \delta_{\eta\eta'} \delta_{\nu\nu'} \delta_{l0} \delta_{l'0}. \quad (2.36)$$

Хотя действие потенциала возмущения V ограничено первым атомом при $l = 0$, однако в силу принятого приближения (вторых ближайших соседей) пространство q должно содержать атомы $l = 0$ и $l = 1$, т. е.

$$Q = \sum_{l=0}^1 |a_l^0\rangle \langle a_l^0|. \quad (2.37)$$

Используя выражения (2.34) — (2.37) в уравнении (2.25), получаем [25, 80]

$$\begin{aligned} \sum_k c_k \delta_{lk} = \sum_k c_k \{ & [G_{l0}(V_{00}\delta_{0k} + V_{01}\delta_{1k}) + \\ & + G_{l1}(V_{10}\delta_{0k} + V_{11}\delta_{1k})] - [G_{l, N-2}(H_{N-2, 0}^0\delta_{0k} + H_{N-2, 1}^0\delta_{1k}) + \\ & + G_{l, N-1}(H_{N-1, 0}^0\delta_{0k} + H_{N-1, 1}^0\delta_{1k})] \}, \end{aligned} \quad (2.38)$$

где

$$\delta_{lk} = \langle a_l^0 | a_k^0 \rangle. \quad (2.39)$$

Остальные матричные элементы имеют вид

$$\Lambda_{lk} = \langle a_l^0 | \Lambda | a_k^0 \rangle, \quad \Lambda = G, V, H^0. \quad (2.40)$$

С помощью выражения (2.36) и определений $\beta_{1,2}$ уравнение (2.38) сводится к

$$c_l = (AG_{l0} - \beta_1 G_{l, N-1} - \beta_2 G_{l, N-2}) c_0 - \beta_2 G_{l, N-1} c_1. \quad (2.41)$$

¹⁾ Поскольку кристалл считается полубесконечным, поверхностным эффектом при $n = N$ мы пренебрегаем.

Учитывая выражения (2.21), (2.28), (2.30), (2.39) и (2.40), имеем

$$G_{l+k} = G_{l, N-k} = N^{-1} \sum_{\theta} e^{i(l+k)\theta} [E - E^0(\theta)]^{-1}. \quad (2.42)$$

При больших N можем записать

$$G_l = (2\pi)^{-1} \int_0^{2\pi} e^{il\theta} [E - E^0(\theta)]^{-1} d\theta, \quad (2.43)$$

согласно (2.30). Таким образом, вместо (2.41) имеем

$$c_l = (AG_l - \beta_1 G_{l+1} - \beta_2 G_{l+2}) c_0 - \beta_2 G_{l+1} c_1 \quad (2.44)$$

для $l \geq 0$. Полагая $l = 1$ и $l = 0$ в (2.44), получаем, что нетривиальные решения для c_0 и c_1 существуют, если

$$\begin{vmatrix} AG_0 - \beta_1 G_1 - \beta_2 G_2 - 1 & -\beta_2 G_1 \\ AG_1 - \beta_1 G_2 - \beta_2 G_3 & -\beta_2 G_2 - 1 \end{vmatrix} = 0, \quad (2.45)$$

что можно было также получить, исходя из (2.27). Уравнение (2.45) определяет энергию E , отвечающую волновой функции $|\psi\rangle$ из (2.35).

1) *Энергия и коэффициенты волновой функции поверхностного состояния.* Перейдем в формуле (2.43) к интегрированию по окружности единичного радиуса в комплексной плоскости

$$G_l = \frac{i}{2\pi\beta_2} \oint \frac{z^{l+1}}{F(z)} dz. \quad (2.46)$$

Здесь, согласно (1.9),

$$F(z) = \frac{2z^2(X^0 - X)}{\rho}. \quad (2.47)$$

С помощью соотношений (2.31) и (2.32) это выражение можно переписать в виде

$$F(z) = z^4 + \rho^{-1}(z^3 + z) - 2X\rho^{-1}z^2 + 1. \quad (2.48)$$

Корни уравнения

$$F(z) = 0 \quad (2.49)$$

равны

$$z_{1,2} = y_1 [1 \pm (1 - y_1^{-2})^{1/2}], \quad z_{3,4} = y_2 [1 \pm (1 - y_2^{-2})^{1/2}], \quad (2.50)$$

причем

$$y_{1,2} = (4\rho)^{-1} \{-1 \pm [1 + 8\rho(X + \rho)]^{1/2}\}. \quad (2.51)$$

Если положить

$$y_{1,2} = \cos \theta_{1,2} \quad (2.52)$$

так, что

$$z_{1,2} = \exp(\pm i\theta_1), \quad z_{3,4} = \exp(\pm i\theta_2), \quad (2.53)$$

то наше рассмотрение [75] станет идентичным подходу Фейрберна [78], который пользовался прямым методом кристаллических орбиталей. Результаты в точности совпадают.

С помощью теоремы о вычетах [63]¹⁾ вместо (2.46) запишем

$$\beta_2 G_l = - \sum_{j=1}^4 R_j, \quad (2.54)$$

где вычеты определяются по формуле

$$R_j = \left[\frac{(z - z_j) z^{l+1}}{\prod_{j=1}^4 (z - z_j)} \right]_{z=z_j} \quad (2.55)$$

Если теперь любой из корней лежит на единичной окружности, то $X = X^0(\theta_{1,2})$ для некоторого действительного значения $\theta_{1,2}$. Однако так как в (2.46) $X \neq X^0(\theta_{1,2})$, то отсюда следует, что величина $\theta_{1,2}$ должна быть комплексной, причем $|y_{1,2}| > 1$ и $|z_{1,3}| > 1$. Таким образом, полюса с вычетами $R_{1,3}$ лежат вне окружности единичного радиуса, и, согласно (2.54) и (2.55), получаем

$$\beta_2 G_l = (z_4 - z_2)^{-1} (1 - z_1 z_3)^{-1} [z_2^l (z_2 - z_1)^{-1} - z_4^l (z_4 - z_3)^{-1}] \quad (2.56)$$

Мы здесь использовали условие

$$z_1 z_2 = z_3 z_4 = 1. \quad (2.57)$$

¹⁾ См. также, например, [318]. — Прим. перев.

Теперь из уравнения (2.45) имеем

$$\alpha = \frac{A}{\beta_2} = \frac{1 + \beta_1 G_1 + 2\beta_2 G_2 + \beta_2^2 (G_2^2 - G_1 G_3)}{\beta_2 G_0 - \beta_2^2 (G_1^2 - G_0 G_2)}, \quad (2.58)$$

что после подстановки (2.56) с учетом равенств

$$\begin{aligned} 1 + \beta_1 G_1 + 2\beta_2 G_2 &= (z_1^3 z_3^3 - z_2 z_4) \Xi, \\ \beta_2^2 (G_2^2 - G_1 G_3) &= z_2 z_4 \Xi, \quad \beta_2 G_0 = (1 - z_1^2 z_3^2) \Xi, \\ \beta_2^2 (G_2^2 - G_1 G_3) &= z_2 z_4 \Xi, \\ \Xi &= (1 - z_1 z_3)^{-2} (z_2 - z_1)^{-1} (z_4 - z_3)^{-1} \end{aligned} \quad (2.59)$$

дает

$$\alpha = -z_1 z_3. \quad (2.60)$$

Несколько утомительные выкладки с использованием соотношений (2.50), (2.51) и (2.60) приводят к выражению для энергии

$$2X = \frac{A^2 + \beta_2^2}{A\beta_1} + \frac{A\beta_1}{(A - \beta_2)^2}, \quad (2.61)$$

а из формулы (2.44) совместно с (2.56) и (2.60) получаем

$$\frac{c_p}{c_0} = \frac{z_4^{p+1} - z_2^{p+1}}{z_4 - z_2}. \quad (2.62)$$

2) *Условия существования поверхностных состояний.* Поскольку полюса $z_{1,3}$ находятся вне окружности единичного радиуса, из (2.60) непосредственно следует, что

$$|\alpha| > 1. \quad (2.63)$$

Если фаза θ_k ($k = 1, 2$) комплексна, то величина y_k может быть либо действительной, либо комплексной в зависимости от вида θ_k . Именно когда $\theta_k = n_k \pi + i\mu_k$, y_k действительно, а когда $\theta_k = \zeta_k + i\mu_k$, то y_k комплексно.

Из (2.51) следует, что y_k действительно, если

$$\delta p(X + p) > -1, \quad (2.64)$$

причем, учитывая (2.51) и (2.52), имеем

$$|y_2| > |y_1| > 1, \quad (2.65)$$

или, возвращаясь к (2.50),

$$|z_3| > |z_4|, \quad |z_3| > |z_1| > |z_2|. \quad (2.66)$$

Переписывая (2.51) в виде

$$2\rho y_k^2 + y_k - (X + \rho) = 0, \quad (2.67)$$

при условии (2.65) имеем

$$|X - \rho| > 1. \quad (2.68)$$

Записав формулу (2.61) с помощью параметров ρ и α

$$2X = \frac{\rho(1 + \alpha^2)}{\alpha} + \frac{\alpha}{\rho(\alpha - 1)^2} \quad (2.69)$$

и воспользовавшись условием (2.68), получаем

$$|\rho|(\alpha - 1)^2 > |\alpha|. \quad (2.70)$$

Условие (2.63) является необходимым условием существования поверхностных состояний; оно становится достаточным, только если величина y_k комплексна. Согласно (2.51), это означает, что должно выполняться неравенство

$$8\rho(X + \rho) < -1, \quad (2.71)$$

так что

$$y_1 = y_2^* \quad (2.72)$$

и в (2.50)

$$z_1 = z_3^*, \quad z_2 = z_4^*, \quad |z_1| > |z_2|. \quad (2.73)$$

Соотношения (2.69) и (2.71) совместно дают

$$2|\rho|(1 + |\alpha|) > |\alpha|^{1/2}, \quad (2.74)$$

или при $|\rho| < 1/4$

$$\alpha < 1 - (8\rho^2)^{-1} [1 + (1 - 16\rho^2)^{1/2}], \quad (2.75)$$

что, очевидно, удовлетворяет требованию (2.63). Если $|\alpha|$ велико, то условие (2.74) приближенно можно записать в виде

$$4\rho^2|\alpha| > 1. \quad (2.76)$$

Это же условие можно получить, считая в (2.75) ρ малой величиной. Сумма корней $y_1, 2$, согласно (2.51), дает

$$2\rho(y_1 + y_2) = -1. \quad (2.77)$$

Однако из (2.50) имеем

$$2y_1 = z_1 + z_1^{-1}, \quad 2y_2 = z_3 + z_3^{-1}. \quad (2.78)$$

В результате вместо (2.77), учитывая (2.60), получаем

$$z_1 + z_3 = \alpha(\rho^{-1} + z_1 + z_3). \quad (2.79)$$

Полагая далее в (2.60) $\alpha < -1$, имеем

$$(z_1 + z_3)^2 > 4z_1z_3 > 4, \quad (2.80)$$

откуда, согласно (2.79),

$$|\rho| > \frac{1}{4}. \quad (2.81)$$

3) *Обсуждение.* На фиг. 9 показаны участки плоскости с координатными осями α и $|\rho|^{-1}$, удовлетворяющие различным условиям существования поверхностных состояний. В приближении сильной связи имеется единственное условие существования поверхностных состояний таммовского типа [1, 2, 14]

$$\left| \frac{A}{\beta_1} \right| = \lim_{\beta_1 \rightarrow 0} |\alpha\rho| > 1. \quad (2.82)$$

У нас ему отвечает условие (2.63), которое представлено на фиг. 9 пунктиром. Как следует из взаимного расположения сплошных и пунктирных линий, при $\alpha > 0$ $\alpha_{\text{ВБС}} > \alpha_{\text{ПСС}}$, а при $\alpha < 0$ $\alpha_{\text{ВБС}} < \alpha_{\text{ПСС}}$ ¹⁾. Таким образом, взаимодействие вторых ближайших соседей затрудняет (облегчает) возникновение поверхностных состояний при $\alpha > 0$ ($\alpha < 0$). Заметим еще, что для малых $|\rho|$ условие (2.70) приближенно имеет вид

$$|\rho(\alpha - 2)| > 1. \quad (2.83)$$

¹⁾ Здесь индекс ВБС означает приближение вторых ближайших соседей, а индекс ПСС — приближение сильной связи. — *Прим. перев.*

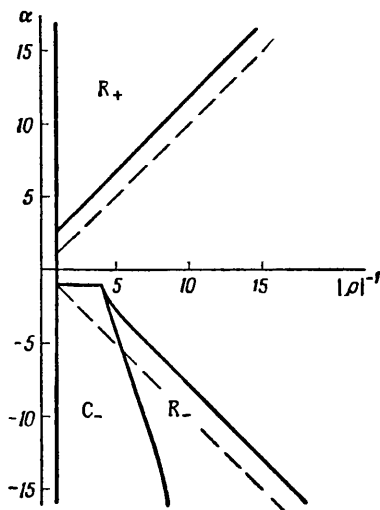
Если y_k действительно (этому отвечают области $R_{\pm}$ на фиг. 9), то соотношение (2.62) можно переписать в виде

$$\left(\frac{c_p}{c_r}\right)^2 = \frac{z_4^{2(p-r)}(1-\omega^{p+1})^2}{(1-\omega^{r+1})^2}, \quad (2.84)$$

где

$$\omega = \frac{z_2}{z_4}. \quad (2.85)$$

Из (2.57) и (2.66) следует, что $|z_4| < 1$, т. е. c_p убывает с увеличением расстояния от поверхности, и элек-



Фиг. 9. Схема, показывающая выполнение условий существования на плоскости $\alpha | \rho |^{-1}$.

Область R_+ (R_-) отвечает действительным значениям y_k при $\alpha > 0$ ($\alpha < 0$); область C_- — комплексным значениям y_k и $\alpha < 0$. Пунктирная линия соответствует приближению сильной связи. Обратите внимание, что $| \rho^{-1} | > 1$, так как $| \beta_2 | > | \beta_1 |$.

трон локализован у поверхности кристалла. Для случая когда y_k комплексно (область C_- на фиг. 9), из (2.53), (2.73) и (2.85) имеем

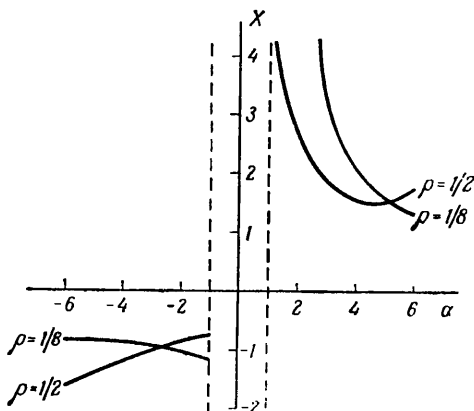
$$\omega = e^{2i\theta}, \quad (2.86)$$

так что вместо формулы (2.84) получаем

$$\left| \frac{c_p}{c_r} \right|^2 = |z_4|^{2(p-r)} \left| \frac{\sin(p+1)\theta}{\sin(r+1)\theta} \right|^2. \quad (2.87)$$

Наличие тригонометрического множителя означает, что на этот раз коэффициенты разложения волновой функции осциллируют, уменьшаясь по величине при удалении от поверхности.

Выражение для энергии поверхностного состояния (2.69) при условии (2.63) проиллюстрировано на фиг. 10

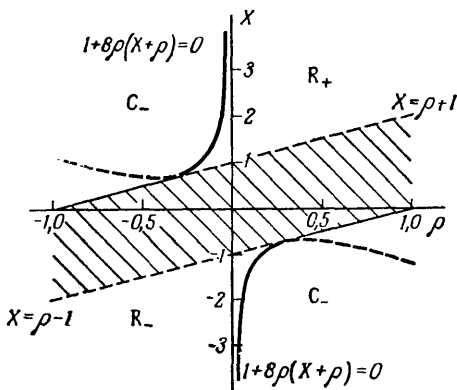


Фиг. 10. Зависимость энергии X от параметра α для $\rho = 1/8$ и $1/2$.

графиком величины X как функции α для $\rho = 1/8$ и $1/2$. Кривые, отвечающие случаям $\rho = -1/8$ и $-1/2$, получаются отражением графиков фиг. 10 относительно оси $X = 0$. Кривая зависимости энергии X от параметра ρ показана на фиг. 11, причем области, в которых выполнены неравенства (2.64), (2.68) и (2.71), помечены в соответствии с обозначениями фиг. 9. Пунктирные линии показывают края зоны объемных состояний; в области между этими линиями не выполнено ни одно из указанных неравенств.

В дополнение к мультипликативной комбинации осциллирующего и затухающего членов в (2.87) при действительных y_k , когда либо θ_1 действительно, а $\theta_2 =$

$= n_2\pi + i\mu_2$ ($\rho > 0$), либо θ_2 действительно, а $\theta_1 = n_1\pi + i\mu_1$ ($\rho < 0$), Фейрберн [78] получил их *аддитивную* комбинацию в выражении для коэффициентов волновой функции. Для удобства классификации мы будем называть эти состояния *псевдоповерхностными* (объемными) [81], если вклад затухающего слагаемого больше (меньше), чем вклад осциллирующего. Энергии этих



Ф и г. 11. Полный энергетический спектр.

Энергии объемных (поверхностных) состояний лежат внутри (вне) области, ограниченной пунктиром. Энергиям псевдосостояний отвечает заштрихованная область.

псевдосостояний, возникающих как результат включения взаимодействий между вторыми ближайшими соседями, лежат, в нарушение условия (2.68), в интервале $\rho - 1 < X < \rho + 1$, т. е. *попадают* [77] в зону объемных состояний (заштрихованная область на фиг. 11). Эти псевдосостояния можно назвать также резонансными (§ 8) ¹⁾.

В приближении сильной связи β_2 стремится к нулю и выражение (2.61) сводится к виду

$$2X_0 = \frac{A}{\beta_1} + \frac{\beta_1}{A}. \quad (2.88)$$

¹⁾ По аналогии с теорией спектра колебаний решетки состояния такого типа можно было бы назвать квазилокальными. См. также работу [319]. — *Прим. перев.*

Следовательно,

$$\Delta X = 2(X - X_0) = \rho [2\sigma^2 + \alpha^{-1}(1 + 3\sigma^2)], \quad (2.89)$$

где

$$\sigma = (\alpha\rho)^{-1}. \quad (2.90)$$

Согласно (2.89), отсюда следует, что

$$|\Delta X|_{\alpha > 0} > |\Delta X|_{\alpha < 0}. \quad (2.91)$$

3. Другие работы, использующие приближение взаимодействия вторых ближайших соседей

Бальдок [74] применил прямой метод кристаллических орбиталей для расчета энергии поверхностных состояний в двумерной одновалентной квадратной металлической решетке. Он указал на тот прогресс, который достигается при учете взаимодействия вторых ближайших соседей. В следующем году Костер и Слэтер [67] рассматривали примесные уровни в линейном кристалле с помощью приближений сильной связи и взаимодействия вторых ближайших соседей. В обобщении этой работы на случай трех измерений была применена техника функций Грина. При этом также учитывались взаимодействия более высоких порядков [82, 83]. Коутецкий [76] рассмотрел поверхностные состояния шоклиевского типа [17], пользуясь резольвентным методом кристаллических орбиталей в одномерной цепочке одинаковых атомов, соединенных связями чередующейся силы. Были учтены резонансные интегралы между ближайшими и вторыми ближайшими соседними атомами. Оказалось, что поверхностные состояния заполнены, если на поверхности разрывается сильная связь. Здесь же были сопоставлены поверхностные состояния таммовского [14] и шоклиевского [17] типов.

Фейрберн [78] допускал, что поверхностное возмущение не ограничено одной лишь первой элементарной ячейкой, но проникает и во вторую. Ему удалось при этом обобщить результаты предыдущих работ [75, 77] и получить дополнительные поверхностные состояния. Кроме того, подобная модель позволяет исследовать влияние взаимодействия вторых ближайших соседей на элек-

тронные процессы хемосорбции чужеродных атомов на кристаллических поверхностях¹⁾). Фейрберн отметил также, что попытка Дэвисона и Дерко [77] упростить свои расчеты в приближении вторых ближайших соседей пренебрежением одним из граничных условий приводит к некоторым ошибочным результатам. Роль граничных условий в приближениях сильной связи и вторых ближайших соседей, их влияние на свойства таммовских поверхностных состояний в полубесконечных и конечных линейных кристаллах были затем детально рассмотрены Дэвисоном и Гриндлеем [79].

¹⁾ Сравните фиг. 4 работы [78] и фиг. 1 работы [1].

МЕТОДЫ КРИСТАЛЛИЧЕСКОГО ПОТЕНЦИАЛА

Будем считать, что потенциал ограниченного кристалла $V(\mathbf{r})$ известен во всем пространстве. Тогда в результате решения уравнения Шредингера получатся все состояния системы, в том числе и поверхностные. В принципе для данного потенциала $V(\mathbf{r})$ задача может быть решена точно и все энергии могут быть вычислены. Однако на практике бывает удобно упростить вид функции $V(\mathbf{r})$ вблизи поверхности. Часто используется подход, заключающийся во введении разрыва потенциала наподобие ступеньки, отделяющей одну область от другой. В предельном случае ступенька, очевидно, превращается в бесконечно высокую стенку, что налагает на волновую функцию условие обращения ее в нуль на поверхности.

Крониг и Пенни [13], изучая квантовомеханическое поведение электрона в кристаллической решетке, представили потенциал кристалла в виде линейной совокупности прямоугольных ям, которую они затем преобразовали в цепочку ям в виде δ -функций так, чтобы площадь каждой ямы осталась неизменной. Саксон и Хатнер [84] применили модельный потенциал Кронига — Пенни в виде δ -функций для построения методов исследования электронных свойств различных одномерных моделей кристаллов, основанных на применении матрицы рассеяния и функций Грина. Кристаллическое поле может быть также приближенно описано с помощью синусоидального потенциала [85—88]. В этом случае уравнение Шредингера эквивалентно уравнению Матье [89], аналитические решения которого хорошо известны [90, 91].

§ 3. Модель Кронига — Пенни

1. Релятивистское рассмотрение

Рассмотрим, в порядке обобщения классических работ Кронига — Пенни [13] и Тамма [14], релятивистскую задачу об объемных и поверхностных состояниях в од-

ноатомном кристалле. После формулировки общего дираковского подхода мы воспользуемся методом матрицы рассеяния [38, 84] для получения закона дисперсии. При этом волновые функции объемных состояний непосредственно получаются из уравнения Дирака как результат применения теоремы Блоха. Поверхностная задача решается с помощью процедуры, применявшейся Таммом.

Релятивистские поправки в расчетах энергетической зонной структуры оказываются довольно значительными в случае, когда кристалл состоит из атомов тяжелых элементов. До сих пор в таких релятивистских расчетах [92—98] уравнение Дирака решалось для так называемого ячеечного (muffin-tin) потенциала. При этом использовалось каноническое преобразование Фолди—Войтхузена [99], чтобы расцепить состояния с положительной и отрицательной энергиями и получить в гамильтониане поправочные релятивистские члены порядка $O(\alpha^2)$, где $\alpha = e^2/\hbar c$ — постоянная тонкой структуры. Одноэлектронная волновая функция кристалла выражается затем в виде линейной комбинации присоединенных (или ортогонализованных) плоских волн, коэффициенты которой находятся вариационным методом.

В случае задачи о кристаллической поверхности решение уравнения Дирака для потенциала ячеек становится особенно трудным из-за сложностей, вносимых граничными условиями. Чтобы сделать решение проблемы более доступным, необходимо обратиться к потенциалам более простого вида, например к выбранному здесь потенциалу Кронига—Пенни [100, 101].

а. Общая теория

1) *Формулировка Дирака.* Движение электрона с массой покоя m_0 в поле линейного потенциала $V(x)$ описывается уравнением Дирака [102]

$$\left(i\hbar \frac{\partial}{\partial t} - i\hbar c \alpha_x \nabla_x + \beta m_0 c^2 - V \right) \psi = 0, \quad (3.1)$$

где c — скорость света, ψ — четырехкомпонентная спинорная волновая функция, зависящая от переменных x

и t ,

$$\psi(x, t) = \begin{bmatrix} \psi_{(1)}(x, t) \\ \psi_{(2)}(x, t) \\ \psi_{(3)}(x, t) \\ \psi_{(4)}(x, t) \end{bmatrix}, \quad (3.2)$$

$$\alpha_x = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \beta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \quad (3.3)$$

Подстановка (3.2) и (3.3) в (3.1) дает систему четырех уравнений:

$$i\hbar\psi_{(1)} - i\hbar c\psi'_{(4)} + m_0 c^2 \psi_{(1)} - V\psi_{(1)} = 0, \quad (3.4a)$$

$$i\hbar\psi_{(2)} - i\hbar c\psi'_{(3)} + m_0 c^2 \psi_{(2)} - V\psi_{(2)} = 0, \quad (3.4б)$$

$$i\hbar\psi_{(3)} - i\hbar c\psi'_{(2)} - m_0 c^2 \psi_{(3)} - V\psi_{(3)} = 0, \quad (3.4в)$$

$$i\hbar\psi_{(4)} - i\hbar c\psi'_{(1)} - m_0 c^2 \psi_{(4)} - V\psi_{(4)} = 0. \quad (3.4г)$$

Можно заметить, что уравнения (3.4a) и (3.4г), а также (3.4б) и (3.4в) образуют пары, которые могут быть записаны в матричной форме в виде

$$i\hbar\psi - i\hbar c\sigma_x\psi' + m_0 c^2\sigma_z\psi - V\psi = 0, \quad (3.5)$$

где σ_x, σ_z — x - и z -компоненты спиновых матриц Паули:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (3.6)$$

а

$$\psi = \begin{bmatrix} \psi_{(1,2)} \\ \psi_{(4,3)} \end{bmatrix} \quad (3.7)$$

теперь обозначает двухкомпонентную спинорную функцию x и t . Если потенциал $V(x)$ не зависит от времени, то пространственные и временные переменные в уравнении (3.5) разделяются путем подстановки

$$\psi(x, t) = \phi(x) e^{-iEt/\hbar}. \quad (3.8)$$

Итак, уравнение (3.5) после подстановки (3.8) сводится к одномерному уравнению Дирака, не зависящему от времени,

$$i\hbar\sigma_x\phi' - m_0c^2\sigma_z\phi = (E - V)\phi, \quad (3.9)$$

решение которого дает стационарные состояния в поле одномерного потенциала $V(x)$.

2) *Решение для постоянного потенциала.* Положив

$$\phi = \begin{bmatrix} \phi_{(1)} \\ \phi_{(2)} \end{bmatrix} \quad (3.10)$$

и пользуясь выражениями (3.6), можно записать уравнение (3.9) в виде

$$i\hbar c\phi'_{(2)} = [(\varepsilon - V) + 2m_0c^2]\phi_{(1)}, \quad (3.11)$$

$$i\hbar c\phi'_{(1)} = (\varepsilon - V)\phi_{(2)}, \quad (3.12)$$

причем

$$\varepsilon = E - m_0c^2. \quad (3.13)$$

Расцепление уравнений (3.11) и (3.12) путем их дифференцирования дает

$$\phi''_{(j)} = -\kappa_v^2\phi_{(j)}, \quad j = 1 \text{ и } 2, \quad (3.14)$$

где

$$\kappa_v^2 = \frac{1}{\hbar^2 c^2} (\varepsilon - V)[(\varepsilon - V) + 2m_0c^2]. \quad (3.15)$$

Решение уравнений (3.14) для постоянного потенциала V дает плоские волны

$$\phi_{(j)} = \alpha_1^{(j)} \exp(i\kappa_v x) + \alpha_2^{(j)} \exp(-i\kappa_v x); \quad (3.16)$$

в результате функцию (3.10) можно записать в виде

$$\phi = A_1 \exp(i\kappa_v x) + A_2 \exp(-i\kappa_v x), \quad (3.17)$$

где

$$A_t = \begin{bmatrix} \alpha_t^{(1)} \\ \alpha_t^{(2)} \end{bmatrix}, \quad t = 1 \text{ и } 2. \quad (3.18)$$

Из (3.12) и (3.16) следует, что

$$\alpha_t^{(1)} = \mp \gamma \alpha_t^{(2)}, \quad (3.19)$$

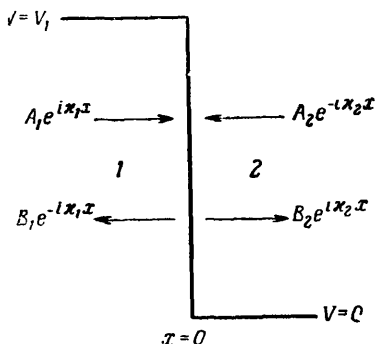
где

$$\gamma = \frac{\varepsilon - V}{\hbar c \kappa_V}. \quad (3.20)$$

Знак «минус» или «плюс» отвечает значениям $t = 1$ и 2 соответственно. Объединяя выражения (3.17)–(3.19), получаем решение

$$\phi = \begin{bmatrix} -\gamma \\ 1 \end{bmatrix} \alpha_1^{(2)} \exp(i\kappa_V x) + \begin{bmatrix} \gamma \\ 1 \end{bmatrix} \alpha_2^{(2)} \exp(-i\kappa_V x). \quad (3.21)$$

3) *Решение для потенциала в виде «ступеньки».* Взаимодействие свободного электрона с потенциалом,



Фиг. 12. Процесс рассеяния на потенциальной «ступеньке».

имеющим ступеньку в точке $x = 0$, может быть рассмотрено как задача рассеяния [38, 84]. Падающие и рассеянные волны (фиг. 12)¹⁾ связаны матрицей рассеяния

$$B = SA, \quad (3.22)$$

где

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad (3.23)$$

¹⁾ В противоположность работе Глассера и Дэвисона [101] здесь принята положительная потенциальная ступенька.

причем

$$A_k = \begin{bmatrix} \alpha_k^{(1)} \\ \alpha_k^{(2)} \end{bmatrix}, \quad B_k = \begin{bmatrix} \beta_k^{(1)} \\ \beta_k^{(2)} \end{bmatrix} \quad (3.24)$$

и

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}. \quad (3.25)$$

Волны по обе стороны от рассеивающего потенциала связаны R -матрицей:

$$\begin{bmatrix} B_2 \\ A_2 \end{bmatrix} = R \begin{bmatrix} A_1 \\ B_1 \end{bmatrix}, \quad (3.26)$$

где

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}. \quad (3.27)$$

Пользуясь соотношениями (3.22) и (3.26), можно выразить элементы одной из матриц S и R через элементы другой:

$$S = R_{22}^{-1} \begin{bmatrix} -R_{21} & 1 \\ R_{11}R_{12} - R_{12}R_{21} & R_{12} \end{bmatrix}, \quad (3.28)$$

$$R = S_{12}^{-1} \begin{bmatrix} S_{12}S_{21} - S_{11}S_{22} & S_{22} \\ -S_{11} & 1 \end{bmatrix}. \quad (3.29)$$

Таким образом, матрицы S и R бесконечно велики, если $R_{22} = 0$ или $S_{12} = 0$ соответственно (см. п. 2в § 3).

В областях, обозначенных номерами 1 и 2 (фиг. 12), не зависящие от времени спиноры могут быть представлены в виде

$$\phi_1 = A_1 \exp(i\kappa_1 x) + B_1 \exp(-i\kappa_1 x), \quad (3.30)$$

$$\phi_2 = B_2 \exp(i\kappa_2 x) + A_2 \exp(-i\kappa_2 x), \quad (3.31)$$

где введены обозначения ¹⁾

$$\kappa_1^2 = \frac{1}{\hbar^2 c^2} (\epsilon - V_1) [(\epsilon - V_1) + 2m_0 c^2], \quad (3.32)$$

$$\kappa_2^2 = \frac{\epsilon (\epsilon + 2m_0 c^2)}{\hbar^2 c^2}. \quad (3.33)$$

¹⁾ См. примечание на предыдущей странице.

Условие сшивания двухкомпонентных спиноров в точке разрыва потенциала при $x = 0$ сводится к условию непрерывности

$$\phi_1(0) = \phi_2(0). \quad (3.34)$$

Проблема выбора условия непрерывности и его роль в анализе релятивистских поверхностных и примесных состояний недавно были подробно исследованы в работах [59, 103, 104]. С помощью выражений (3.19), (3.24), (3.30), (3.31) и условия (3.34) получаем

$$\Gamma(\beta_1^{(2)} - \alpha_1^{(2)}) = \alpha_2^{(2)} - \beta_2^{(2)}, \quad (3.35)$$

$$\alpha_1^{(2)} + \beta_1^{(2)} = \alpha_2^{(2)} + \beta_2^{(2)}, \quad (3.36)$$

где

$$\Gamma = \frac{v_1}{v_2}, \quad v_k = \frac{e - V}{\hbar c x_k}. \quad (3.37)$$

Из уравнений (3.35) и (3.36) находим

$$\begin{aligned} \beta_1^{(2)} &= \frac{2\alpha_2^{(2)} - (1 - \Gamma)\alpha_1^{(2)}}{1 + \Gamma}, \\ \beta_2^{(2)} &= \frac{2\Gamma\alpha_1^{(2)} + (1 - \Gamma)\alpha_2^{(2)}}{1 + \Gamma}. \end{aligned} \quad (3.38)$$

Если теперь сравнить эти соотношения с уравнением (3.22), записанным для компонент, то сразу получаем

$$\begin{aligned} S_{11} &= -\frac{1 - \Gamma}{1 + \Gamma}, & S_{12} &= \frac{2}{1 + \Gamma}, \\ S_{21} &= \Gamma S_{12}, & S_{22} &= -S_{11}. \end{aligned} \quad (3.39)$$

Подстановка (3.39) в (3.29) дает выражение для матрицы R :

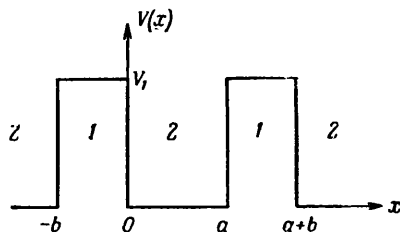
$$R_0 = \begin{bmatrix} v_+ & v_- \\ v_- & v_+ \end{bmatrix}, \quad v_{\pm} = \frac{1 \pm \Gamma}{2}, \quad (3.40)$$

отвечающее рассеянию в точке $x = 0$.

б. Объемные свойства

1) *Закон дисперсии.* В модели Кронига — Пенни [13] кристаллический потенциал представляется последовательностью прямоугольных ям (фиг. 13). При движении

вдоль оси x преобразование спинора, отвечающее смещению на расстояние d без прохождения через скачок



Ф и г. 13. Потенциал Кронига — Пенни.

потенциала, может быть описано с помощью следующей трансляционной матрицы [84]:

$$T_d = \begin{bmatrix} e^{i\kappa d} & 0 \\ 0 & e^{-i\kappa d} \end{bmatrix}. \quad (3.41)$$

Таким образом, смещение на расстояние, равное периоду $a + b$, сопровождается умножением спинора на матрицу

$$T = T_b R_a T_a R_0. \quad (3.42)$$

Рассматривая соответственно области 1 и 2 и полагая

$$\chi_1 = \kappa_1 b, \quad \chi_2 = \kappa_2 a, \quad (3.43)$$

получаем с помощью выражений (3.40) и (3.41) матричные элементы матрицы (3.42)

$$\begin{aligned} T_{11} &= \exp(i\chi_1) [2\Gamma \cos \chi_2 + i(1 + \Gamma^2) \sin \chi_2] / 2\Gamma, \\ T_{12} &= i(1 - \Gamma^2) \exp(i\chi_1) \sin \chi_2 / 2\Gamma, \\ T_{21} &= T_{12}^*, \quad T_{22} = T_{11}^*, \end{aligned} \quad (3.44)$$

где звездочка означает комплексное сопряжение. Матрица T является унитарной, т. е.

$$\det T = 1. \quad (3.45)$$

Теорема Блоха — Флоке [36, 38, 84, 105] требует, чтобы волновое число θ удовлетворяло равенству

$$\cos \theta a = \frac{1}{2} \operatorname{tr} T. \quad (3.46)$$

С помощью выражений (3.44) и (3.46) получаем

$$\cos \theta a = \cos \chi_1 \cos \chi_2 - \Xi \sin \chi_1 \sin \chi_2, \quad (3.47)$$

где

$$\Xi = \frac{\Gamma + \Gamma^{-1}}{2}. \quad (3.48)$$

Если преобразовать потенциал Кронига — Пенни в последовательность δ -образных потенциалов, перейдя к пределу так, чтобы выполнялось условие

$$\lim_{\substack{V_1 \rightarrow \infty \\ b \rightarrow 0}} V_1 b = p, \quad (3.49)$$

то соответствующие предельные значения величин χ_1 и γ_1 будут равны

$$\chi_1 = -\frac{p}{\hbar c}, \quad \gamma_1 = 1. \quad (3.50)$$

Подставляя (3.37), (3.43), (3.48) и (3.50) в равенство (3.47) и принимая, что $p \ll \hbar c$, получаем *трансцендентное релятивистское соотношение Кронига — Пенни* для $\kappa_2 a$:

$$\cos \theta a = \cos \kappa_2 a + \frac{mp}{\hbar^2 \kappa_2} \sin \kappa_2 a, \quad (3.51)$$

$$m = \frac{E}{c^2} = m_0 \left(1 + \frac{v^2}{2c^2} \right), \quad (3.52)$$

причем m является электронной массой с учетом релятивистской поправки первого порядка. В нерелятивистском пределе, когда $c \rightarrow \infty$, из (3.33) и (3.52) имеем

$$\kappa_2 = \kappa = \frac{(2m_0 e)^{1/2}}{\hbar}, \quad m = m_0, \quad (3.53)$$

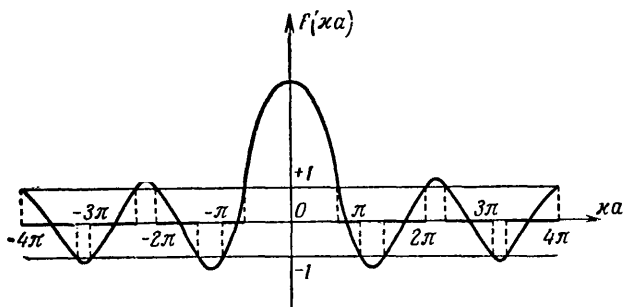
что сводит уравнение (3.51) к обычному закону дисперсии модели Кронига — Пенни

$$f(\kappa a) = \cos \theta a = \cos \kappa a + \frac{m_0 p}{\hbar^2 \kappa} \sin \kappa a. \quad (3.54)$$

Сравнивая теперь уравнения (3.51) и (3.54), можно видеть, что релятивистская формулировка Дирака приводит в точности к такому же виду закона дисперсии Кронига — Пенни, что и в нерелятивистском случае. Реля-

тивистские эффекты проявляются лишь в величинах волнового вектора κ_2 и массы электрона.

График функции $f(\kappa a)$ для значения $P = m_0 a p / \hbar^2 = 3\pi/2$ представлен на фиг. 14. Поскольку $|\cos \theta a| \leq 1$, а экстремумы функции $f(\kappa a)$ лежат вне полосы ± 1 , то отрезки кривой, попадающие при изменении θ в этот



Фиг. 14. График функции $f(\kappa a)$ при $P = 3\pi/2$. Жирными линиями отмечены разрешенные энергетические зоны.

интервал, будучи спроектированы на ось κa , дадут разрешенные зоны (жирные линии), разделенные запрещенными зонами. Если величина P положительна, то верхние края зон определяются формулами [106]

$$\kappa a = n\pi, \quad n = 0, \pm 1, \dots \quad (3.55)$$

Для $P < 0$ соотношение (3.55) определяет нижние края зон.

Теперь с помощью выражений (3.33) и (3.53) получаем

$$\kappa_2 = \kappa \left(1 + \frac{\eta^2 \chi^2}{2} \right), \quad (3.56)$$

где

$$\eta = \frac{\hbar}{2m_0 a c}, \quad \chi = \kappa a. \quad (3.57)$$

Аналогично из (3.13), (3.52) и (3.53) имеем

$$m = m_0 (1 + 2\eta^2 \chi^2). \quad (3.58)$$

При переходе к атомным единицам

$$m_0 = \frac{1}{2}, \quad \hbar = 1, \quad e^2 = 2 \quad (3.59)$$

единицей длины будет боровский радиус $a_0 = 0,529 \text{ \AA}$, единицей энергии — ридберг ($1 \text{ Ry} = 13,6 \text{ эВ}$), и вместо выражения (3.57) и уравнения (3.51) получаем

$$\eta = (ac)^{-1}, \quad \eta\chi = \frac{x}{c}, \quad (3.60)$$

$$g(\chi) = \cos \theta a = \cos \chi \left(1 + \frac{\eta^2 \chi^2}{2} \right) + \\ + \frac{pa(1 + 2\eta^2 \chi^2)}{2(1 + \eta^2 \chi^2/2)} \sin \chi \left(1 + \frac{\eta^2 \chi^2}{2} \right). \quad (3.61)$$

Подобно формуле (3.55) верхние границы зон разрешенных значений χ с учетом релятивистских поправок первого порядка определяются условием

$$\kappa_2 a = n\pi. \quad (3.62)$$

Подставляя в левую часть выражение (3.56), приходим к кубическому уравнению

$$\eta^2 \chi^3 + 2\chi - 2n\pi = 0, \quad (3.63)$$

решение которого [107] имеет вид

$$\chi = \eta^{-1} [(r + s)^{1/3} + (r - s)^{1/3}], \quad (3.64)$$

причем

$$r^2 = s^2 + \frac{8}{27}, \quad s = \eta n\pi. \quad (3.65)$$

Положения нижних границ зон можно определить, если приближенно с помощью метода Ньютона [107] найти корни уравнения

$$g(\chi) + (-1)^n = 0. \quad (3.66)$$

В качестве примера рассмотрим решетку с параметром $a = 6a_0 \approx 3 \text{ \AA}$. В единицах (3.59) получаем $c = 274$, так как постоянная тонкой структуры

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137}.$$

Согласно выражению (3.60), $\eta = 6 \cdot 10^{-4}$. В соответствии с условием $pa \ll c$ выберем значение p так, чтобы

$$pa = 5 \cdot 10^{-3}.$$

В табл. 1 показаны сдвиги центров зон, изменения их ширины и ширины запрещенных зон в зависимости от номера зоны n для принятых значений параметров. Так как по ширине разрешенные зоны близки к своему наибольшему значению π , запрещенные зоны очень малы. Фактически при $n \geq 5$ запрещенные зоны так малы, что релятивистские поправки для ширины зоны больше, чем их ширина [96]. Для низких (высоких) энергий ширина нерелятивистской зоны увеличивается (стремится к константе), тогда как с учетом релятивистских поправок первого порядка она сначала растет до максимального

Таблица 1

Результаты расчета энергетического спектра

для $a = 6a_0 > 3\text{\AA}$, $pa = 5 \cdot 10^{-3}$, $\eta = 6 \cdot 10^{-4}$ *

n	Центр зоны		Ширина зоны		Ширина запрещенной зоны	
	χ	χ_2	$\Delta\chi$	$\Delta\chi_2$	$\Delta\chi$	$\Delta\chi_2$
1	1,60614	1,60614	3,07090	3,07089	0,07070	0,07070
2	4,71318	4,71316	3,14000	3,13996	0,00159	0,00159
3	7,85438	7,85428	3,14080	3,14069	0,00080	0,00080
4	10,99584	10,99559	3,14106	3,14086	0,00053	0,00053
5	14,13737	14,13684	3,14119	3,14085	0,00040	0,00040
6	17,27892	17,27797	3,14127	3,14077	0,00032	0,00032
7	20,42048	20,41893	3,14133	3,14062	0,00027	0,00027
8	23,56206	23,55967	3,14136	3,14042	0,00023	0,00023
9	26,70364	26,70018	3,14139	3,14018	0,00020	0,00020
10	29,84522	29,84040	3,14142	3,13991	0,00018	0,00018
15	45,55315	45,53609	3,14148	3,13796	0,00011	0,00011
20	61,26110	61,21972	3,14151	3,13516	0,00008	0,00008
25	76,96905	76,88714	3,14153	3,13153	0,00007	0,00007
30	92,67701	92,53427	3,14154	3,12708	0,00006	0,00006

* Если $V(\chi)$ положительно, первая запрещенная зона возникает при $n=1$.

значения при $n = 4$, а затем убывает с ростом n . Напротив, на ширине запрещенной зоны релятивистские эффекты не сказываются, и после первоначального уменьшения она стремится к постоянному значению при высоких энергиях. Таким образом, для больших n релятивистские эффекты вызывают *сжатие* спектра в целом [97], причем это сжатие состоит в уменьшении ширины разрешенных зон, оставляя запрещенные зоны неизменными. Для больших значений pa (порядка 4) запрещенные зоны с учетом релятивистских эффектов первого порядка шире, чем в нерелятивистском случае, особенно для высоких энергий. Варьируя параметр η (pa) при постоянном pa (η), находим, что при учете релятивистских поправок первого порядка с уменьшением (увеличением) η (pa) наибольшей шириной обладает зона с большим номером n . Несмотря на простой вид потенциала Кронига—Пенни, который был здесь использован, полученные результаты качественно согласуются с экспериментом и релятивистскими расчетами по методу присоединенных плоских волн [97].

2) *Выражение для спиноров.* Согласно теореме Блоха [36—39], имеем

$$\phi_k(x) = e^{i\theta x} u_k(x), \quad (3.67)$$

где $u_k(x)$ — периодическая спинорная функция с периодом, равным периоду решетки, а индекс k обозначает соответствующую область потенциала. После подстановки функции (3.67) в уравнение (3.14) и решения соответствующего дифференциального уравнения второго порядка получаем

$$u_k(x) = A_k \exp[i(\kappa_k - \theta)x] + B_k \exp[-i(\kappa_k + \theta)x], \quad (3.68)$$

где коэффициенты A_k и B_k определены согласно (3.24). Условия непрерывности для $u_k(x)$

$$u_1(0) = u_2(0), \quad u_1(-b) = u_2(a) \quad (3.69)$$

дают

$$A_1 + B_1 = A_2 + B_2, \quad (3.70)$$

$$\begin{aligned} A_1 \exp[-i(\kappa_1 - \theta)b] + B_1 \exp[i(\kappa_1 + \theta)b] = \\ = A_2 \exp[i(\kappa_2 - \theta)a] + B_2 \exp[-i(\kappa_2 + \theta)a], \end{aligned} \quad (3.71)$$

откуда, согласно определениям (3.43), (3.49) и (3.50) и в предположении, что $p \ll \hbar c$, получаем

$$A_2 \{1 - \exp[i(\kappa_2 - \theta)a]\} = B_2 \{\exp[-i(\kappa_2 + \theta)a] - 1\}. \quad (3.72)$$

Вторые компоненты этого матричного уравнения приводят к выражению

$$\lambda = \frac{\beta_2^{(2)}}{\alpha_2^{(2)}} = \frac{1 - \exp[i(\kappa_2 - \theta)a]}{\exp[-i(\kappa_2 + \theta)a] - 1}. \quad (3.73)$$

Возвращаясь к формулам (3.67) и (3.68), находим спинорные решения для объемных состояний:

$$\phi_2(x) = A_2 \exp(i\kappa_2 x) + B_2 \exp(-i\kappa_2 x) \quad (3.74)$$

или, пользуясь выражениями (3.19) и (3.73),

$$\phi_2(x) = \alpha_2^{(2)} \left\{ \begin{bmatrix} -\gamma_2 \\ 1 \end{bmatrix} \exp(i\kappa_2 x) + \lambda \begin{bmatrix} \gamma_2 \\ 1 \end{bmatrix} \exp(-i\kappa_2 x) \right\}. \quad (3.75)$$

в. Таммовская задача о поверхности

В кристаллах, построенных из атомов тяжелых элементов, сжатие спектра, вызванное релятивистскими эффектами, существенно сдвигает положения краев зон, особенно в области высоких энергий. Отсюда вытекает необходимость учета релятивистских эффектов, поскольку положения границ зон весьма важны при расчетах поверхностных состояний.

1) *Выражения для энергии.* Решение уравнения Дирака в области вакуума (фиг. 15) имеет вид [108]

$$\phi_0(x) = \begin{bmatrix} \gamma_0 \\ 1 \end{bmatrix} \beta_0^{(2)} \exp(l_0 x), \quad (3.76)$$

где $l_0 = -i\kappa_0$ — действительная положительная величина. Условие непрерывности на поверхности при $x=0$

$$\phi_0(0) = \phi_2(0) \quad (3.77)$$

после подстановки решений (3.75) и (3.76) дает

$$\gamma_0 \beta_0^{(2)} + \alpha_2^{(2)} \gamma_2 (1 - \lambda) = 0, \quad \beta_0^{(2)} - \alpha_2^{(2)} (1 + \lambda) = 0. \quad (3.78)$$

Нетривиальное решение существует, если

$$\gamma_0 (1 + \lambda) + \gamma_2 (1 - \lambda) = 0. \quad (3.79)$$

С помощью выражения (3.73) перепишем равенство (3.79) в виде

$$e^{i\theta a} = \cos \chi_2 - i\Gamma_0 \sin \chi_2, \quad (3.80)$$

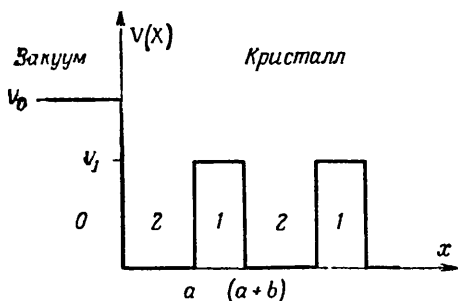
где

$$\Gamma_0 = \frac{\gamma_0}{\gamma_2}. \quad (3.81)$$

Чтобы исследовать вопрос о поверхностных состояниях, допустим, что θ принимает комплексные значения [14]:

$$\theta = \frac{n\pi}{a} + i\mu, \quad (3.82)$$

где μ — действительная положительная величина и n — целое число [ср. выражения (1.32) и (3.82)]. Релятиви-



Фиг. 15. Потенциал полубесконечного линейного кристалла (потенциал поверхности $V_0 > e$). Из работы [108].

стский закон дисперсии Кронига — Пенни (3.51) теперь принимает вид

$$(-1)^n \operatorname{ch} \mu a = \cos \chi_2 + \frac{P'}{\chi_2} \sin \chi_2, \quad (3.83)$$

причем

$$P' = \frac{ma\mu}{\hbar^2}. \quad (3.84)$$

Вычитая уравнение (3.80) из (3.83) и используя (3.82), находим

$$(-1)^n \operatorname{sh} \mu a = \left(\frac{P'}{\chi_2} + i\Gamma_0 \right) \sin \chi_2. \quad (3.85)$$

Воспользовавшись тождеством

$$\operatorname{ch}^2 \mu a - \operatorname{sh}^2 \mu a = 1, \quad (3.86)$$

из (3.83) и (3.85) получаем выражение для энергии поверхностного состояния

$$\chi_2 \operatorname{ctg} \chi_2 = \frac{\chi_2^2}{2P'} \left\{ 1 - \frac{(\epsilon - V_0)}{\epsilon \chi_0^2} \left[\chi_2^2 \frac{(\epsilon - V_0)}{\epsilon} - 2i\chi_0 P' \right] \right\}. \quad (3.87)$$

В первом приближении его можно представить в виде

$$\begin{aligned} \chi \left(1 + \frac{\eta^2 \chi^2}{2} \right) \operatorname{ctg} \chi \left(1 + \frac{\eta^2 \chi^2}{2} \right) = \\ = Q [1 + \eta^2 (\Omega_-^2 - \chi^2)] - \Omega_- \left(1 + \frac{\eta^2 \Omega_+^2}{2} \right), \end{aligned} \quad (3.88)$$

где

$$\Omega_{\pm}^2 = a^2 q^2 \pm \chi^2, \quad Q = \frac{a^2 q^2}{2P'}, \quad q^2 = \frac{2m_0 V_0}{\hbar^2}. \quad (3.89)$$

В нерелятивистском пределе, когда $c \rightarrow \infty$ и $\eta \rightarrow 0$, уравнение (3.88) сводится к знакомому уравнению Тамма [14]

$$\chi \operatorname{ctg} \chi = Q - \Omega_-. \quad (3.90)$$

2) *Условия существования.* Учитывая определения (3.37) и (3.81), вместо уравнения (3.85) находим

$$(-1)^n \operatorname{sh} \mu a = \frac{\sin \chi_2}{\chi_2} \left[P' + \frac{i\chi_2^2 (\epsilon - V_0)}{\epsilon \chi_0} \right], \quad (3.91)$$

где

$$\chi_2 = \chi + \delta, \quad \delta = \frac{\eta^2 \chi^3}{2}, \quad (3.92)$$

что можно получить с помощью определений (3.15), (3.56) и (3.57). В нерелятивистском случае $(n+1)$ -й запрещенной зоне отвечает интервал $n\pi < \chi < (n+1)\pi$. Однако с учетом релятивистского поправочного члена первого порядка δ , согласно (3.92), $(n+1)$ -я запрещенная зона будет теперь определяться условиями

$$(n\pi - \delta) < \chi < [(n+1)\pi - \delta], \quad n = 0, 1, \dots \quad (3.93)$$

Для любой запрещенной зоны нерелятивистский предел уравнения (3.91) дает [109]

$$\operatorname{sgn}\left(\frac{\sin \chi}{\chi}\right) = (-1)^n, \quad (3.94)$$

в результате

$$P > \Omega_-, \quad (3.95)$$

поскольку для $\mu a > 0$ имеем $\operatorname{sh} \mu a > 0$. Если неравенство (3.95) заменить равенством, то его корни будут определять нижние границы зон у поверхности кристалла¹⁾. Верхние их границы определяются соотношением (3.55). В релятивистском приближении первого порядка равенство (3.94) выполняется в интервале энергий (а) $[(n+1)\pi - \delta] > \chi > n\pi$, однако в другой области (б) $n\pi > \chi > n\pi - \delta$ (где $n \neq 0$, так как $\chi > 0$) равенство (3.94) следует заменить таким:

$$\operatorname{sgn}\left(\frac{\sin \chi}{\chi}\right) = (-1)^{n+1}. \quad (3.96)$$

Следовательно, в интервалах (а) и (б), учитывая (3.15), (3.81), (3.84) и (3.91), соответственно получаем

$$P - \Omega_- \geq \frac{1}{2} \eta^2 (\Omega_- \Omega_+^2 - 4P\chi^2). \quad (3.97)$$

В нерелятивистском пределе, когда $\eta \rightarrow 0$, верхнее неравенство переходит в таммовское условие существования (3.95), поэтому поверхностные состояния, удовлетворяющие этому неравенству, мы будем называть *релятивистскими таммовскими состояниями*. С другой стороны, нижнее неравенство в нерелятивистском пределе означает нарушение таммовского условия существования. Поверхностные состояния, удовлетворяющие нижнему неравенству, возникают, таким образом, исключительно как результат релятивистской дираковской формулировки, и по этой причине мы будем называть их *дираковскими поверхностными состояниями*.

3) *Классификация поверхностных состояний по энергии*. Если превратить неравенства (3.97) в уравнения,

¹⁾ Если в нерелятивистском случае значения χ отвечают нижним краям зон у поверхности кристалла, в релятивистском приближении первого порядка они соответствуют поверхностным уровням энергии, лежащим внутри запрещенной зоны.

то значения χ , которые им удовлетворяют, будут отвечать нижним краям зон у поверхности в релятивистском приближении первого порядка. Положения поверхностных уровней энергии относительно различных запрещенных зон получаются в этом приближении в результате решения уравнения (3.88). Для существования поверхностных состояний требуется выполнение одного из неравенств (3.97), что, собственно, и обуславливает возможность найти соответствующее решение (3.88). Однако эта оговорка относится только к релятивистским таммовским состояниям в первой запрещенной зоне [верхнее неравенство (3.97)]. Для дираковских поверхностных состояний с $n \geq 1$ (нижнее неравенство) решения (3.88) в запрещенных зонах с номерами $n \geq 1$ существуют всегда.

В нерелятивистском пределе для значений χ , отвечающих нижним границам зон, левая часть (L) неравенства (3.97) обращается в нуль, так что

$$P = \Omega_- \quad (3.98)$$

Правая часть (R) (3.97) имеет вид

$$R = \frac{\eta^2 P (\Omega_-^2 - 2\chi^2)}{2} \quad (3.99)$$

Если $R < 0$, верхнее из неравенств (3.97) всегда удовлетворено, и

$$\frac{V_0}{3} < \varepsilon < V_0, \quad (3.100)$$

где мы учли соотношения (3.53), (3.57) и (3.89). Если $L = 0$ и $R < 0$, то нижнее неравенство (3.97) никогда не выполняется. Если, напротив, $R > 0$, то верхнее неравенство не может быть удовлетворено, а нижнее удовлетворено всегда. Это отвечает интервалу энергий

$$0 < \varepsilon < \frac{V_0}{3}. \quad (3.101)$$

Пусть теперь нерелятивистское условие для таммовских состояний (3.95) не выполнено, т. е. $L < 0$. Если $R < 0$, то можно одновременно удовлетворить обоим неравенствам (3.97) везде, кроме интервала энергий (3.100). В случае $R > 0$ верхнее неравенство всегда нарушено, однако нижнее всегда выполнено, что отвечает интер-

валу энергий (3.101). С другой стороны, когда $L > 0$, выполнено условие (3.95). Взяв $R > 0$, получим, что оба неравенства могут выполняться по отдельности, обеспечивая выполнение (3.101). Если $R < 0$ и (3.100) выполнено, то нижнее (верхнее) неравенство всегда нарушено (выполнено).

На фиг. 16 представлена схема классификации поверхностных состояний по величине их энергии в нереля-

	$R < 0$	$D_s T_s$	T_a	T_a
$V_0/3$	$R > 0$	D_a	D_a	$D_s T_s$
	$L < 0$	$L = 0$	$L > 0$	

Фиг. 16. Классификация поверхностных состояний в соответствии со значениями энергии ϵ (из работы [108]).

Буквами R (L) обозначена правая (левая) часть неравенства (3.97), D и T — дираковские и релятивистские таммовские состояния. Индексы s или a указывают, что поверхностное состояние «может» существовать или «всегда» существует соответственно.

лятивистском пределе ϵ и значениям параметров P , q и a , которые определяют, будет ли величина L больше, равна или меньше нуля. Пусть $L > 0$, т. е. выполнено нерелятивистское таммовское условие. Тогда в случае $R > 0$ могут присутствовать либо релятивистские таммовские состояния, либо дираковские поверхностные состояния, но когда $R < 0$, всегда имеются только релятивистские таммовские состояния. Таким образом, аналогично нерелятивистскому случаю состояния таммовского типа могут появиться во всем энергетическом интервале. Напротив, дираковские поверхностные состояния могут возникать в нижней части энергетического спектра, если только там нет релятивистских таммовских состояний. Далее, когда $R < 0$ ($R > 0$), реля-

тивистские таммовские состояния возникают легче (труднее), чем таммовские состояния в нерелятивистском случае. Если $L = 0$, что отвечает (в нерелятивистском пределе) нижним границам зон у поверхности кристалла, в полном энергетическом спектре всегда присутствует релятивистское поверхностное состояние того или иного типа. Для этого частного значения L прямая $\varepsilon = V_0/3$ отделяет область энергий, отвечающих релятивистским таммовским состояниям и дираковским поверхностным состояниям. Разумеется, в нерелятивистском пределе для такого значения L поверхностные состояния не возникают. Когда $L < 0$, нерелятивистское таммовское условие (3.95) нарушено, однако дираковские поверхностные состояния могут существовать во всей доступной области энергий. Резюмируя, можно сказать, что, согласно схеме на фиг. 16, дираковские поверхностные состояния (релятивистские таммовские состояния) всегда существуют, если $R > 0$ ($R < 0$) и $L \leq 0$ ($L \geq 0$), и что в области $D_c T_c$ требуется, чтобы величины L и R были положительными или отрицательными одновременно.

Объемные [341] и поверхностные [342] состояния в двухатомных кристаллах были недавно рассмотрены с помощью релятивистской модели Кронига — Пенни. Вместе с тем вопрос о существовании дираковских поверхностных состояний стал предметом некоторой дискуссии. Первоначально Сабраманиан и Бхагват [320], а затем Фейрберн, Глассер и Стенслица [343] исследовали причины появления этих состояний. Дэвисон [344] и, в особенности, Стенслица [345] на основании тщательного изучения точных условий существования подтвердили наличие поверхностных состояний этого типа.

2. Нерелятивистское рассмотрение

В п. 1 § 3 мы рассмотрели объемные свойства кристалла, исходя из решения уравнения Дирака с помощью метода матрицы рассеяния. Теперь же мы применим технику функций Грина [84] для решения уравнения Шредингера. Однако, переходя к проблеме поверхности, мы снова обратимся к методу матрицы рассеяния.

а. Общая теория

Уравнение Шредингера в случае одномерной системы можно записать в виде

$$\phi''(x) + \kappa^2 \phi(x) = -\left(\frac{2m_0}{\hbar^2}\right) V(x) \phi(x), \quad (3.102)$$

где величина κ определяется согласно (3.53) и $V(x) > 0$. Первый шаг при решении уравнения (3.102) состоит в замене правой части δ -функцией Дирака и нахождении решения более простого уравнения

$$G''(x, \tau) + \kappa^2 G(x, \tau) = -\delta(x - \tau), \quad (3.103)$$

которое и определяет функцию Грина $G(x, \tau)$ [102, 110].

Уравнение (3.103) решается с помощью фурье-преобразования [63, 110]. Пусть $g(k)$ — фурье-образ функции $G(x, \tau)$, т. е.

$$G(x, \tau) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} g(k) \exp[ik(x - \tau)] dk. \quad (3.104)$$

Подставляя выражение (3.104) в уравнение (3.103), имеем

$$\int_{-\infty}^{\infty} [(k^2 - \kappa^2) g(k) - (2\pi)^{-1/2}] \exp[ik(x - \tau)] dk = 0,$$

где x и τ могут быть любыми, так что часть подынтегральной функции, заключенная в квадратные скобки, должна обратиться в нуль. Следовательно,

$$g(k) = (2\pi)^{-1/2} (k^2 - \kappa^2)^{-1}. \quad (3.105)$$

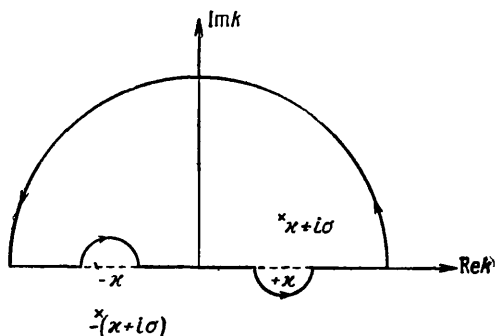
Вместо (3.104) теперь получаем

$$G(x, \tau) = (2\pi)^{-1} \int_{-\infty}^{\infty} (k^2 - \kappa^2)^{-1} \exp[ik(x - \tau)] dk. \quad (3.106)$$

Подынтегральная функция в выражении (3.106) расходится, когда $k \rightarrow \pm \kappa$. Чтобы обойти эту трудность, перепишем функцию $G(x, \tau)$ в виде

$$G(x, \tau) = (2\pi)^{-1} \lim_{\sigma \rightarrow 0} \int_{-\infty}^{\infty} \frac{\exp[ik(x - \tau)] dk}{[k^2 - (\kappa + i\sigma)^2]}. \quad (3.107)$$

Если взять $\sigma > 0$, то полюса в подынтегральном выражении формулы (3.107) отвечают значениям $k = \pm(\kappa + i\sigma)$, и контур интегрирования в комплексной k -плоскости (фиг. 17) называют *выходящим*. Интеграл



Фиг. 17 «Выходящий» контур интегрирования для вычисления $G(x, \tau)$ в комплексной плоскости k .

по полуокружности большого радиуса отвечает комплексным значениям k , очень большим по абсолютной величине, поэтому экспонента в подынтегральной функции стремится к нулю. С помощью теоремы о вычетах имеем

$$G(x, \tau) = (2\pi)^{-1} \lim_{\sigma \rightarrow 0} [2\pi i \times \text{Res}(\kappa + i\sigma)]. \quad (3.108)$$

Вычет в точке $(\kappa + i\sigma)$ получается, если положить $k = (\kappa + i\sigma + \nu)$ и найти коэффициент при ν^{-1} в разложении подынтегральной функции в ряд Лорана, т. е.

$$\frac{\exp[i(\kappa + i\sigma)(x - \tau)]}{2(\kappa + i\sigma)}. \quad (3.109)$$

Таким образом, вместо (3.108) будем иметь

$$G(x, \tau) = \frac{i \exp(i\kappa|x - \tau|)}{2\kappa}. \quad (3.110)$$

Решение (3.110) уравнения (3.103) можно теперь использовать для решения уравнения (3.102). Умножая уравнение (3.102) и (3.103) на G и ϕ соответственно,

интегрируя по τ и затем вычитая одно из получившихся уравнений из другого, найдем

$$\begin{aligned} \int_{-\infty}^{\infty} [\phi(\tau) G''(\tau) - G(\tau) \phi''(\tau)] d\tau = \\ = \frac{2m_0}{\hbar^2} \int_{-\infty}^{\infty} V(\tau) G(\tau) \phi(\tau) d\tau - \int_{-\infty}^{\infty} \phi(\tau) \delta(x - \tau) d\tau. \end{aligned} \quad (3.111)$$

Поскольку левая часть равенства (3.111), согласно теореме Грина [63, 110], равна нулю, имеем

$$\phi(x) = \frac{2m_0}{\hbar^2} \int_{-\infty}^{\infty} V(\tau) G(x, \tau) \phi(\tau) d\tau. \quad (3.112)$$

Здесь учтено свойство δ -функции [40, 69]:

$$\phi(x) = \int_{-\infty}^{\infty} \phi(\tau) \delta(x - \tau) d\tau. \quad (3.113)$$

Уравнение (3.112) показывает, что с помощью функции Грина можно свести дифференциальное уравнение Шредингера (3.102) к некоторому интегральному уравнению.

б. Объемные решения

Полагая параметр решетки равным a , заменим интеграл в (3.112) суммой интегралов по каждой элементарной ячейке, т. е.

$$\phi(x) = \frac{2m_0}{\hbar^2} \sum_{n=-\infty}^{\infty} \int_{na}^{(n+1)a} V(\tau) G(x, \tau) \phi(\tau) d\tau. \quad (3.114)$$

Учитывая периодичность кристаллического потенциала, имеем

$$V(\tau + na) = V(\tau). \quad (3.115)$$

Тогда из выражения (3.110) получаем

$$G(x, \tau + na) = \frac{i \exp(i\kappa |x - \tau - na|)}{2\kappa}. \quad (3.116)$$

Теорема Блоха (3.67), кроме того, дает

$$\phi(\tau + na) = \exp(i\theta na) \phi(\tau). \quad (3.117)$$

В результате уравнение (3.114) можно записать в виде

$$\phi(x) = \frac{im_0}{\hbar^2 \kappa} \sum_{n=-\infty}^{\infty} e^{i\theta na} \int_0^a \exp(i\kappa|x - \tau - na|) \phi(\tau) V(\tau) d\tau \quad (3.118)$$

или

$$\phi(x) = \frac{im_0}{\hbar^2 \kappa} \int_0^a F(x, \tau) \phi(\tau) V(\tau) d\tau, \quad (3.119)$$

где

$$F(x, \tau) = \left[\sum_{-\infty}^{[t/a]} + \sum_{[t/a]+1}^{\infty} \right] \exp(i\theta na) \exp(i\kappa|t - na|), \quad (3.120)$$

причем $t = x - \tau$, а величина $[t/a]$ обозначает целую часть переменной t/a . Положив $t' = t - a[t/a]$ и просуммировав геометрическую прогрессию, после несколько утомительных выкладок получаем вместо (3.120)

$$F(x, \tau) = -i \exp\{i\theta a[t/a]\} \times \\ \times \left[\frac{e^{i\theta a \sin \kappa t'} - \sin \kappa(t' - a)}{\cos \kappa a - \cos \theta a} \right]. \quad (3.121)$$

Теперь уравнение (3.119) имеет вид

$$\phi(x) = \frac{m_0}{\hbar^2 \kappa} \int_0^a \exp\{i\theta a[t/a]\} \times \\ \times \left[\frac{e^{i\theta a \sin \kappa t'} - \sin \kappa(t' - a)}{\cos \kappa a - \cos \theta a} \right] \phi(\tau) V(\tau) d\tau. \quad (3.122)$$

Потенциал линейного одноатомного кристалла можно аппроксимировать цепочкой δ -функций Дирака. Если выбрать за начало координат один из узлов решетки, то полный потенциал системы можно записать в виде

$$\frac{2m_0}{\hbar^2} V(x) = -aU \sum_{n=-\infty}^{\infty} \delta(x - na). \quad (3.123)$$

Параметр $U < 0$ описывает интенсивность δ -образных потенциальных ям в каждом узле решетки. Подставив (3.123) в (3.122) и воспользовавшись равенством (3.113), для первой элементарной ячейки, где $\tau = 0$ и $t' = x'$, получаем

$$\phi(x) = \frac{aU}{2\kappa} \phi(0) \exp[i\theta(x - x')] \times \\ \times \left[\frac{e^{i\theta a} \sin \kappa x' - \sin \kappa(x' - a)}{\cos \theta a - \cos \kappa a} \right]. \quad (3.124)$$

Полагая $x' = x = 0$, мы сразу возвращаемся к закону дисперсии Кронига — Пенни [ср. с (3.54)]

$$\cos \theta a = \cos \kappa a + \frac{aU}{2\kappa} \sin \kappa a. \quad (3.125)$$

в. Поверхностные решения

1) *Волновые функции.* Рассмотрим полубесконечный линейный кристалл (a — параметр решетки). Представим потенциал кристалла (фиг. 18) выражением

$$\frac{2m_0}{\hbar^2} V(x) = \begin{cases} -aU \sum_{n=0}^{\infty} \delta(x - na), & x > 0, \\ -V_0, & V_0 < 0, \quad x < 0. \end{cases} \quad (3.126)$$

Внутри кристалла ($x > 0$) волновая функция имеет общий вид [111]

$$\phi_1(x) = \{A_1 [e^{-i\theta a} \sin \kappa x - \sin \kappa(x - a)] + \\ + B_1 [e^{i\theta a} \sin \kappa x - \sin \kappa(x - a)]\} \frac{1}{\sin \kappa a} \quad (3.127)$$

в соответствии с (3.124) и (3.125). В области вакуума вне кристалла ($x < 0$) общим решением уравнения (3.102) служит функция

$$\phi_0(x) = A_0 \exp(i\kappa_0 x) + B_0 \exp(-i\kappa_0 x), \quad (3.128)$$

где

$$\kappa_0^2 = \kappa^2 + V_0. \quad (3.129)$$

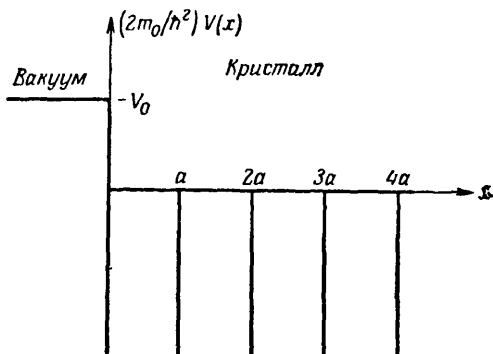
2) *Граничные условия и S-матрица.* Условие непрерывности волновой функции в точке $x = 0$ требует [ср. с (3.34)] следующего равенства:

$$\phi_0(0) = \phi_1(0). \quad (3.130)$$

Используя равенство (3.113) и интегрируя уравнение Шредингера в интервале $0_- < x < 0_+$, получаем¹⁾

$$\phi'(0_+) - \phi'(0_-) = aU\phi(0). \quad (3.131)$$

После подстановки волновых функций (3.127) и (3.128)



Фиг. 18. Потенциал полубесконечного линейного кристалла, состоящий из δ -функций Дирака.

в равенства (3.130) и (3.131) имеем соответственно

$$A_0 + B_0 = A_1 + B_1, \quad (3.132)$$

$$\frac{\kappa}{\sin \kappa a} [A_1(e^{-i\theta a} - \cos \kappa a) + B_1(e^{i\theta a} - \cos \kappa a)] - i\kappa_0(A_0 - B_0) = aU(A_0 + B_0). \quad (3.133)$$

Если переписать уравнения (3.132) и (3.133) в матричном виде (3.22), то элементы S-матрицы, согласно (3.28), дают [111]

$$R_{22} = e^{i\theta a} - \cos \kappa a - \frac{aU - i\kappa_0}{\kappa} \sin \kappa a. \quad (3.134)$$

¹⁾ Необходимость введения граничного условия (3.131) вместо более привычного условия $\phi'(0_+) = \phi'(0_-)$ связано с конечной «силой» δ -образной потенциальной ямы.

3) *Уравнения для поверхностных состояний.* Как уже отмечалось в связи с соотношением (3.28), элементы матрицы S из (3.22) стремятся к бесконечности, когда $R_{22} = 0$, при этом амплитуды B становятся бесконечно большими по сравнению с амплитудами A . Тогда если для выходящей волны θ действительно, то поверхность действует как источник электронов, что с физической точки зрения бессмысленно. С другой стороны, если θ определяется согласно (3.82), то выходящие волны затухают, и электроны локализованы у поверхности. Итак, условия существования поверхностных состояний имеют вид

$$R_{22} = 0, \quad \theta = \frac{n\pi}{a} + i\mu, \quad (3.135)$$

где μ — действительная положительная величина.

4) *Уровни энергии.* С помощью выражений (3.125), (3.134) и (3.135) получаем

$$(-1)^n \operatorname{ch} \mu a = \cos \kappa a + \frac{aU}{2\kappa} \sin \kappa a, \quad (3.136)$$

$$(-1)^n e^{-\mu a} = \cos \kappa a + \frac{aU - i\kappa_0}{\kappa} \sin \kappa a. \quad (3.137)$$

Исключая из этих уравнений $\cos \kappa a$ [ср. с (3.85)], имеем

$$(-1)^n \operatorname{sh} \mu a = \frac{i\kappa_0 - aU/2}{\kappa} \sin \kappa a. \quad (3.138)$$

Тождество (3.86) совместно с уравнениями (3.136) и (3.138) приводит к уравнению для энергии поверхностного состояния [112]

$$\chi \operatorname{ctg} \chi + i\chi_0 + \frac{V_0}{U} = 0, \quad (3.139)$$

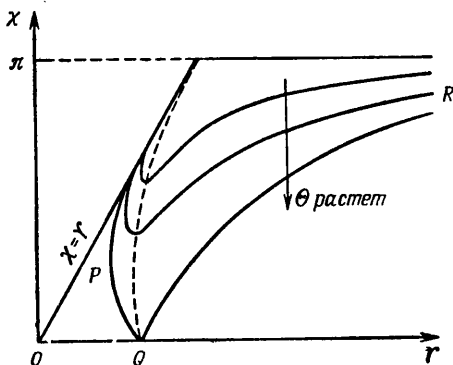
где вновь введено обозначение (3.57). При этом, чтобы энергия поверхностного состояния оказалась действительной, необходимо, чтобы величина χ_0 была чисто мнимой. Далее, поскольку $\exp(-i\kappa_0 x)$ описывает волну, экспоненциально затухающую в отрицательном направлении оси x , то должно выполняться неравенство $i\kappa_0 < 0$. Таким образом, учитывая (3.129) и вводя параметры

$$\Upsilon^2 = -a^2 V_0, \quad \Theta = -a^2 U, \quad (3.140)$$

получаем вместо уравнения (3.139) [ср. с (3.90)]

$$\chi \operatorname{ctg} \chi - (\Upsilon^2 - \chi^2)^{1/2} + \frac{\Upsilon^2}{\Theta} = 0. \quad (3.141)$$

Графические решения этого уравнения схематически показаны на фиг. 19. В плоскости $\chi\Upsilon$ запрещенная зона ограничена асимптотой $\chi = \pi$ и прямой $\chi = \Upsilon$, отделяющей область, в которой решения (3.141) всегда действительны (т. е. при $\chi \leq \Upsilon$). Минимальным значениям переменных Υ и χ отвечают точки $P(\Upsilon_p, \chi_p)$ и $Q(\Upsilon_q, \chi_q)$ соответственно. Геометрическое место точек Q показано на фиг. 19 пунктирной линией, оно опреде-



Фиг. 19. Решения уравнения (3.141), полученные графически.

ляется уравнением

$$\Theta = 2(\Upsilon_q^2 - \chi_q^2)^{1/2}, \quad (3.142)$$

которое в свою очередь получается дифференцированием уравнения (3.141) по Υ при условии, что в точке (Υ_q, χ_q) $d\chi/d\Upsilon = 0$.

Сравнение фиг. 19 с энергетическим спектром модели Кронига — Пенни для отрицательного потенциала¹⁾ показывает, что асимптота $\chi = \pi$ отвечает дну второй

¹⁾ Фаризо [112], взяв, по-видимому по ошибке, положительный потенциал в законе дисперсии Кронига — Пенни, поменял в результате местами верхние и нижние границы каждой зоны.

зоны, тогда как точка Q принадлежит верхней границе первой. Тот же результат можно получить с помощью хорошо известного уравнения [106] для верхних границ зон, а именно

$$\chi \operatorname{ctg} \frac{\chi}{2} = \frac{\Theta}{2}. \quad (3.143)$$

Это уравнение, будучи подставлено в (3.141), вновь приводит к выражению (3.142) для точки Q . Заметим также, что, когда в уравнении (3.142) $\Theta=4$, $\Upsilon_q=2$ и $\chi_q=0$, первая зона пропадает.

5) *Условия существования.* Согласно закону дисперсии Кронига — Пенни (3.125), имеем

$$\operatorname{sgn} \left(\frac{\sin \kappa a}{\kappa} \right) = (-1)^{n+1} \quad (3.144)$$

для $(n-1)\pi < \kappa a < n\pi$, где $n = 1, 2, \dots$ обозначает номер запрещенной зоны. Поэтому, так как $\mu > 0$, уравнение (3.138) удовлетворяется, только когда

$$i\kappa_0 - \frac{aU}{2} < 0. \quad (3.145)$$

Это неравенство, если использовать (3.129), (3.140) и условие $i\kappa_0 < 0$, дает

$$\Theta < 2(\Upsilon^2 - \chi^2)^{1/2}. \quad (3.146)$$

Итак, поверхностные состояния существуют только для тех решений (3.141), которые удовлетворяют неравенству (3.146).

Легко показать [109], что условие существования (3.146) удовлетворено тогда и только тогда, когда $\Upsilon > \Upsilon_q$. Таким образом, на фиг. 19 поверхностным состояниям отвечают отрезки кривых, лежащие справа от пунктирной линии, поскольку только здесь удовлетворено условие (3.146). Оставшимся частям кривых соответствуют значения $\mu < 0$, которые определяют волны с амплитудами, нарастающими в глубь кристалла. Каждой точке ветви кривой QR соответствует одно поверхностное состояние. Отсюда, в согласии с результатом Тамма [14], для каждого данного набора значений параметров Θ , Υ и χ , удовлетворяющих условию (3.146), может су-

ществовать только одно¹⁾ поверхностное состояние. При нарушении условия (3.146) поверхностные состояния не существуют. Как было показано Фаризо [112] (см. также примечание на стр. 75), поведение кривых QR таково, что хотя поверхностные состояния могут исчезнуть, слившись с краем нижней зоны в точках Q , однако ни одно из них не в состоянии достичь дна следующей зоны из-за асимптотического поведения кривых QR в окрестности R .

г. Другие работы

1) *Метод S-матрицы*. Кроме своей работы, посвященной поверхностным состояниям в кристаллах с чистыми поверхностями, Аэртс [111] изучал, пользуясь методом матрицы рассеяния, поверхностные состояния в кристаллах, поверхность которых загрязнена примесью. Аэртс показал, что если потенциал примеси слабее (сильнее), чем потенциал атомов в объеме, то электрон связан с примесным атомом слабее (сильнее), так что примесь можно рассматривать как донорную (акцепторную); поверхностный уровень при этом смещается в сторону зоны проводимости (валентной зоны). Подобным же образом Аэртсу удалось классифицировать поверхностную проводимость как проводимость n (донорного) или p (акцепторного) типов.

После рассмотрения одноатомных кристаллов Аэртс [47] обратился к двухатомным кристаллам с чистыми поверхностями, которые мы детально обсуждали в п. 2 § 1. И вновь главной отличительной чертой системы, в элементарной ячейке которой находятся два атома, является возникновение дополнительной запрещенной зоны. В последней из статей этой серии [113] Аэртс проанализировал проблему шоклиевских поверхностных состояний, постулируя симметричный поверхностный потенциал. Рассматривая адсорбцию посторонних атомов на поверхности, он получил, что поверхностные состояния могут возникнуть, если потенциал примеси меньше,

¹⁾ Это до некоторой степени противоречит результату Аэртса [111], который, пренебрегая условием (3.146), нашел, что максимальное возможное число поверхностных состояний равно *двум*.

чем потенциал атомов в объеме. Кроме того, Аэртс рассчитал энергетические уровни состояний, локализованных у чистой поверхности контакта двух полубесконечных линейных кристаллов, составленных из различных атомов и обладающих разными периодами решетки. Оказалось, что «контактные» уровни возникают в запрещенной зоне, общей для обоих кристаллов. Гош и Ласкар [114] обобщили эту работу, рассмотрев контакты, загрязненные примесями.

2) *Метод функций Грина.* До настоящего времени направление, использующее метод функций Грина, развивалось в основном усилиями Фаризо¹⁾. Полученные им данные по теории поверхностных состояний в полубесконечных [112] и конечных [115] чистых кристаллах²⁾ находятся в согласии с вышеизложенными результатами. Быть может, одной из самых интересных работ Фаризо является та, в которой рассмотрен энергетический спектр аморфных веществ [116]. Используя метод цепных дробей, он показал, что такие вещества, подобно кристаллам, обладают зонной энергетической структурой. Эта структура появляется исключительно как следствие достаточно малого межатомного расстояния и не вызвана какой бы то ни было атомной периодичностью³⁾. Было найдено, что у системы, состоящей из аморфного вещества, окруженного кристаллической оболочкой из того же вещества, в запрещенной зоне присутствуют дополнительные локальные уровни энергии. Эта работа легла в основу последующей статьи [56] о подповерхностных состояниях в деформированных кристаллах (п. 3 § 1), поскольку деформированная часть кристалла, созданная распространением поверхностного возмущения, может рассматриваться как аморфная область.

Аналогично работе Аэртса [47] по бинарным кристаллам типа АВ Фаризо [117] рассмотрел поверхностные состояния в соединениях типа $A^{III}B^V$, имея в виду прежде всего структуру цинковой обманки (или сфалерита),

¹⁾ В последние годы метод функций Грина широко применяется при рассмотрении задачи о поверхностных состояниях [321—323]. — *Прим. перев.*

²⁾ См. также [324]. — *Прим. перев.*

³⁾ См. также обзорную статью [325]. — *Прим. перев.*

причем были взяты одномерные модели для направлений [100] и [111]. Результаты в случае направления [100] совпали с результатами Аэртса [47]. Однако для направления [111], вдоль которого атомы расположены в последовательности двойных слоев (расстояния между ними относятся к толщине слоев, как 1:3), наличие поверхностных состояний в большой степени зависит от того, какого сорта связь разорвана на поверхности. Согласно результатам Коутецкого [118], поверхностные состояния в первой запрещенной зоне отвечают разрыву на поверхности сильной связи. Следует, однако, заметить, что вновь образованная поверхность, конечно, более стабильна, когда нарушена слабая связь. Отсюда следует, что поверхностные состояния должны в таком случае исчезать. Для второй запрещенной зоны уровни энергии в случае разрыва слабой связи отвечают поверхностным потенциальным барьерам более низким, чем те, которые получаются при разрыве сильной связи.

3) *Влияние приложенного поля.* Метод сшивания волновой функции был использован для исследования влияния полей у поверхности на таммовские состояния в рамках модели Кронига — Пенни [346, 347] и задачи Матье [348], которая рассматривается в следующем параграфе. Оказалось, что наличие внешних полей может как скрывать, так и вызывать появление таммовских состояний. Внешние поля могут, кроме того, смещать положение поверхностных энергетических уровней относительно запрещенных зон. Подобные расчеты чрезвычайно важны при изучении туннелирования из поверхностных состояний [329, 349].

§ 4. Задача Матье

Простой моделью кристаллического потенциала, несколько более реалистичной для расчетов поверхностных состояний, чем последовательность δ -функций, служит потенциал в виде синусоиды. Его амплитуда определяет величину кристаллического потенциала в одноатомном и степень ионности связи в двухатомном кристалле. На поверхности кристалл оканчивается потенциальной ступенькой произвольной высоты. Граница может распо-

лагаться в любой точке последней элементарной ячейки; в частных случаях это может быть точка минимума или максимума. Так как уравнение Шредингера с синусоидальным потенциалом приводится к хорошо известному уравнению Матье [91, 119], многие результаты можно получить довольно легко. Статц [21] первым предложил использовать задачу Матье в расчетах поверхностных состояний. Недавно Левин [34] выполнил соответствующие расчеты для случая произвольного положения границы. В нашем последующем рассмотрении мы будем придерживаться такого плана: 1) изучим свойства уравнения Матье; 2) введем поверхность кристалла; 3) исследуем условие непрерывности; 4) рассмотрим конечный кристалл и учтем гибридизацию всех состояний и, наконец, 5) сравним полученные результаты для поверхностных состояний с результатами других моделей. Мы увидим, что таммовские, шоклиевские и ионные поверхностные состояния могут быть успешно сопоставлены в рамках модели Матье.

1. Общая теория

Уравнение Шредингера для одномерного кристалла с синусоидальным потенциалом амплитуды V_2 и постоянным потенциалом V_1 имеет вид

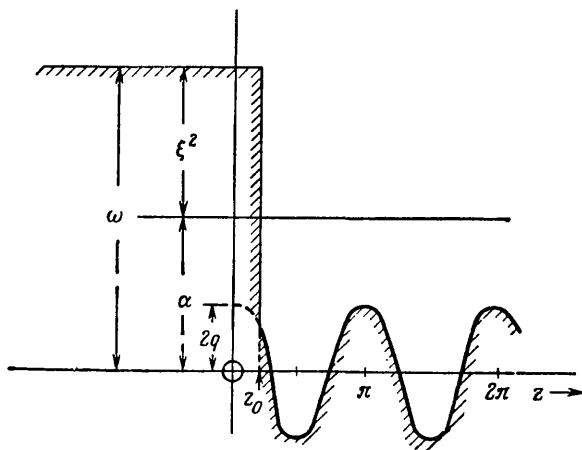
$$\frac{\hbar^2}{2m_0} \psi''(x) + \left(E - V_1 - V_2 \cos \frac{2\pi x}{a} \right) \psi(x) = 0. \quad (4.1)$$

Потенциал имеет период a и по определению достигает максимального значения (отталкивание) в точке $x = 0$. (Сдвиг начала координат будет рассмотрен позднее.) Для удобства запишем уравнение в безразмерном виде

$$\psi''(z) + (\alpha - 2q \cos 2z) \psi(z) = 0. \quad (4.2)$$

Здесь $z = \pi x/a$, α — безразмерная энергия, а параметр q определяет отношение амплитуды потенциальной энергии к энергии кинетической. Нормированный кристаллический потенциал показан на фиг. 20. Уравнение (4.2) является уравнением Матье, записанным в стандартной форме, используемой Мак-Лахланом [91]. (Янке и Эмде [119] применяют несколько иные обозначения.)

В предельном случае, когда $q = 0$, кристаллический потенциал становится постоянным (приближение свободных электронов). Тогда ненормированные решения (4.2) типа стоячих волн отвечают значениям $\alpha = m^2$, где $m = 1, 2, 3, \dots$, причем $\psi = +\cos z, +\sin z, +\cos 2z, +\sin 2z, \dots$. (Для удобства всегда берется знак «плюс».) Имеется также решение для $\alpha = m = 0$, а именно $\psi = 1$.



Фиг. 20. Поверхностная область кристалла с синусоидальным потенциалом.

Если $q \neq 0$, то энергия α и волновая функция ψ должны зависеть от q , что можно записать, используя разложение в ряд. Возьмем, например, решение, которое в пределе $q = 0$ имеет вид $\psi = \cos z$ и $\alpha = 1$ ¹⁾. Соответствующие разложения запишутся как

$$\alpha = 1 + \alpha_1 q + \alpha_2 q^2 + \alpha_3 q^3 + \dots, \quad (4.3)$$

$$\psi = \cos z + q C_1(z) + q^2 C_2(z) + q^3 C_3(z) + \dots \quad (4.4)$$

Подставляя (4.3) и (4.4) в уравнение (4.2) и приравнявая нулю коэффициенты при одинаковых степенях q ,

¹⁾ Значению $\alpha = 1$ отвечает еще и другое решение, которое в пределе $q = 0$ переходит в $\psi = \sin z$.

имеем

$$q^0: \cos z - \cos z = 0, \quad (4.5)$$

$$q^1: C_1'' + C_1 - \cos 3z + (\alpha_1 - 1) \cos z = 0, \quad (4.6)$$

$$q^2: C_2'' + C_2 - \alpha_1 C_1 - 2C_1 \cos 2z + \alpha_2 \cos z = 0. \quad (4.7)$$

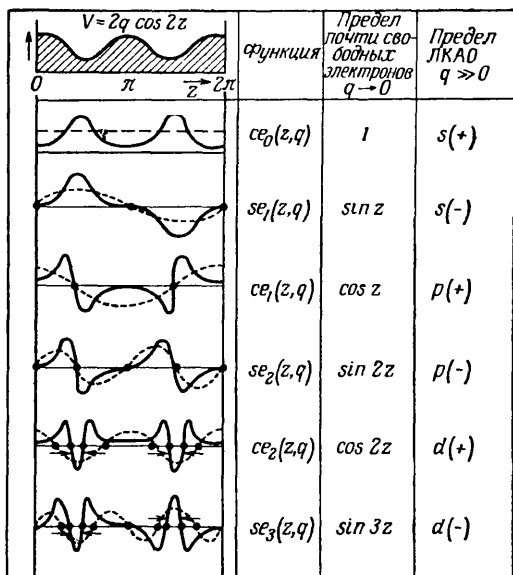
⋮

Эти уравнения могут быть решены по порядку, если использовать условия симметрии и произвести последовательное интегрирование [91]. Например, условия симметрии требуют, чтобы $\alpha_1 = 1$ в уравнении (4.6) и $\alpha_2 = -1/8$ в (4.7). Аналогичным способом можно вычислить все коэффициенты разложений энергии $\alpha_1, \alpha_2, \alpha_3, \dots$ и волновой функции $C_1, C_2, C_3, \dots$. В результате собственное значение α и собственная функция ψ в (4.3) и (4.4) могут быть получены с любой требуемой точностью. К счастью, α и ψ уже *протабулированы* и *построены* как функции q и m [91, 119]. Наиболее важные для нас свойства ψ и α представлены на фиг. 21 и 22 соответственно.

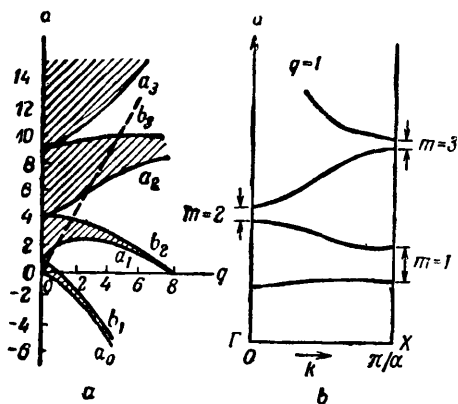
Если $q \neq 0$, то собственные функции типа стоячих волн уравнения Матье (4.2) по определению называются *косинус-функциями* Матье $se_m(z, q)$ и *синус-функциями* Матье $se_m(z, q)^1$. Их поведение качественно показано на фиг. 21. В пределе свободных электронов потенциал решетки постоянен и $q = 0$. Тогда по определению $se_m(z, q) \rightarrow \cos mz$ и $se_m(z, q) \rightarrow \sin mz$. Эти функции представляют собой волновые функции свободного электрона; они нанесены на фиг. 21 пунктирными линиями. Предел почти свободных электронов определяется условием $q \approx 0$, так что $se_m(z, q) \approx \cos mz$ и $se_m(z, q) \approx \sin mz$. С ростом q потенциальные ямы притяжения с центрами в точках $z = \pi/2$ (а также $z = 3\pi/2$ и т. д.) углубляются. Это вызывает смещение максимумов и узлов собственных функций в направлении к точке $z = \pi/2$, т. е. туда, где потенциальная энергия меньше²⁾. При очень больших значениях q собственные функции становятся больше похожими на

¹⁾ В отечественной литературе и те и другие называются периодическими функциями Матье. — *Прим. перев.*

²⁾ Больше по абсолютной величине. — *Прим. перев.*



Фиг. 21. Некоторые свойства функций Матье.
Знак (+) обозначает связывающие, а знак (-) — антисвязывающие орбитали.



Фиг. 22. а — энергетические зоны как функции амплитуды синусоиды; б — энергетические зоны как функции волнового вектора для $q = 1$.

функции метода ЛКАО. Они показаны схематически на фиг. 21 сплошными линиями. Эти функции напоминают s , p , d и т. д., связывающие и антисвязывающие атомные орбитали. Говоря конкретнее, функции $se_0(z, q)$ и $se_1(z, q)$ становятся функциями s -типа (не имеют узлов вблизи $z = \pi/2$), функции $se_1(z, q)$ и $se_2(z, q)$ — функциями p -типа (один узел при $z = \pi/2$), функции $se_2(z, q)$ и $se_3(z, q)$ — d -типа (два узла около $z = \pi/2$) и т. д. Здесь индекс m служит для нумерации соответствующих запрещенных зон. Таким образом, функции Матье дают представление о том, как происходит непрерывный переход между предельными случаями почти свободных электронов и ЛКАО.

Собственные значения, отвечающие стоячим волнам, $\alpha = a_m$ или $\alpha = b_m$ для произвольных значений q — известны; они показаны на фиг. 22, *а* сплошными линиями. Заштрихованные области диаграммы соответствуют разрешенным состояниям типа блоховских волн $e^{\pm i\mu z}\phi(\pm z, q)$. Незаштрихованные участки — запрещенные зоны; волновые функции здесь имеют вид $e^{\mp i\mu z}\phi(\pm z, q)$. При рассмотрении фиг. 22, *а* можно заметить, что, когда $q = 0$, запрещенные зоны отсутствуют; это — предел свободных электронов. С ростом q вырождение снимается, причем сильнее всего это проявляется в первой запрещенной зоне ($m = 1$), где периоды потенциала и волновой функции оказываются совпадающими. В пределе больших q разрешенные зоны становятся все уже (все более напоминают атомные уровни), а их энергии понижаются, следуя увеличению потенциала притяжения вблизи $z = \pi/2$.

На фиг. 22, *б* схематически показана, кроме того, зависимость энергии α от волнового вектора k для значения $q = 1$. Дно первой разрешенной зоны есть состояние a_0 . Первая запрещенная зона находится на грани X зоны Бриллюэна, она ограничена состояниями b_1 и a_1 . Вторая запрещенная зона расположена в центре Γ зоны Бриллюэна и ограничена состояниями b_2 и a_2 . Запрещенные зоны на грани зоны Бриллюэна отвечают нечетным значениям m , а запрещенные зоны в центре зоны Бриллюэна — четным. Так как соединения CdS , ZnS и многие другие кристаллы типа $A^{II}B^{VI}$ и $A^{III}B^{IV}$

имеют основную запрещенную зону в центре Γ трехмерной зоны Бриллюэна¹⁾, то подходящим аналогом этой запрещенной зоны в одномерной задаче Матье может быть запрещенная зона с $m = 2$. В дальнейшем, однако, имеются в виду все запрещенные зоны, а не только самая первая, с $m = 1$.

На этом мы закончим краткий обзор свойств уравнения Матье и объемных состояний в поле синусоидального потенциала. Собственные функции и собственные значения, полученные ранее для произвольных значений q , будут теперь служить основой при рассмотрении различных типов поверхностных состояний и их энергий. По сути дела, поверхностные состояния строятся из описанных выше состояний, отвечающих краям зон. Эти состояния должны быть аналитически продолжены с использованием комплексного волнового вектора.

2. Поверхность

Рассмотрим модель поверхности, схематически изображенную на фиг. 20. Плоскости поверхности отвечает координата z_0 , которая в частном случае может быть точкой максимума ($z_0 = 0$) или минимума ($z_0 = \pi/2$) потенциала в элементарной ячейке. На нашей фигуре положение границы выбрано так, что $z_0 = \pi/8$. Ионные и ковалентные аналоги поверхностей в точках $z_0 = 0, \pi/2, \pi/4$, и $3\pi/4$ будут рассмотрены позже. Пусть, кроме того, кристалл бесконечен в области $z \geq z_0$; следствия из предположения о большой, но конечной толщине кристаллической пластины тривиальны и будут обсуждены в дальнейшем.

Внутри кристалла ($z \geq z_0$) потенциал имеет вид синусоиды. Волновая функция разрешенного поверхностного состояния удовлетворяет внутри кристалла уравнению Матье (4.2) и имеет вид

$$\psi_i(z, q) = \phi(z, q) \exp[-\mu(z - z_0)]. \quad (4.8)$$

Ей отвечает собственное значение α . Здесь по определению μ — положительно, и функция $\psi_i(z, q) \rightarrow 0$, когда $z \rightarrow \infty$.

¹⁾ См., например, [120].

Вне кристалла ($z \leq z_0$) потенциал испытывает скачок величины ω (в безразмерных единицах) по отношению к его среднему значению во внутренней области. Волновую функцию вне кристалла можно записать в форме

$$\psi_0(z, q) = \text{const} \times \exp[-\xi(z_0 - z)] \quad (4.9)$$

с тем же собственным значением α ($\alpha = \omega - \xi^2$). Здесь ξ по определению положительно, и $\psi_0(z) \rightarrow 0$, когда $z \rightarrow -\infty$. Условия непрерывности волновой функции и ее первой производной на поверхности $z = z_0$ дают

$$\frac{\phi(z_0, q)}{\phi'(z_0, q)} = (\xi + \mu)^{-1}. \quad (4.10)$$

Во многих случаях величина, стоящая в правой части, настолько мала, что ее можно рассматривать как возмущение. В пределе бесконечно высокой ступеньки потенциала ($\xi \rightarrow \infty$) условие (4.10) сводится к простому требованию обращения волновой функции в нуль: $\phi(z_0, q) = 0$.

В самом общем виде функция $\phi(z, q)$ будет линейной комбинацией («гибридом») всех объемных стоячих волн $\text{se}_r(z, q)$ и $\text{se}_r(z, q)$ для всех номеров запрещенных зон r . Этот общий случай будет рассмотрен ниже. Превосходную аппроксимацию функции $\phi(z, q)$ можно дать, заметив, что внутри m -й запрещенной зоны поверхностное состояние является суперпозицией в основном функций $\text{se}_m(z, q)$ и $\text{se}_m(z, q)$, которые отвечают границам этой запрещенной зоны. Точнее говоря, функция $\text{se}_m(z, q)$ отвечает собственному значению b_m , лежащему на дне запрещенной зоны, а функция $\text{se}_m(z, q)$ — собственному значению a_m , расположенному на вершине этой запрещенной зоны. (Особый случай $m = 0$ будет разобран позже.) Таким образом, волновая функция поверхностного состояния (4.8) берется в следующем виде (здесь опущен нормировочный множитель):

$$\psi_l(z, q) = [\lambda \text{se}_m(z, q) - \text{se}_m(z, q)] e^{-\mu z}, \quad (4.11)$$

где член в скобках есть $\phi(z, q)$, μ — положительно (по определению), а λ — параметр гибридизации. Это приближение пригодно в тех случаях, когда запрещенная

зона достаточно узка по сравнению с ширинами разрешенных зон ($q < \alpha/2$, как показано пунктиром на фиг. 22, а).

В дополнение к волновым функциям разрешенных поверхностных состояний полезно рассмотреть волновую функцию $\psi_i^*(z, q)$ *запрещенных* поверхностных состояний, которая тоже удовлетворяет уравнению Матье. Она получается заменой в выражении (4.11) $z = -z$, что дает

$$\psi_i^*(z, q) = [\lambda \operatorname{se}_m(z, q) + \operatorname{se}_m(z, q)] e^{\mu z}, \quad (4.12)$$

где μ , как и раньше, — положительная величина. Эта функция экспоненциально *нарастает*, когда $z \rightarrow \infty$, и соответствующее состояние должно *выталкиваться* в случае полубесконечного кристалла.

Параметр гибридизации λ можно выразить через объемные параметры q , m и α , если непосредственно подставить функцию (4.11) в уравнение Матье (4.2), что, согласно методу Мак-Лахлана [34, 91], дает

$$\phi'' - 2\mu\phi' + \mu^2\phi + (\alpha - 2q \cos 2z)\phi = 0. \quad (4.13)$$

Комбинируя (4.11) и (4.13) и учитывая значения интегралов ортогональности [34, 91, 119]

$$\int_0^{2\pi} \operatorname{ce}_m^2(z, q) dz = \int_0^{2\pi} \operatorname{se}_m^2(z, q) dz = \pi, \quad (4.14a)$$

$$\int_0^{2\pi} \operatorname{ce}_m(z, q) \operatorname{se}_m(z, q) dz = 0, \quad (4.14б)$$

$$\int_0^{2\pi} \operatorname{ce}_m(z, q) \operatorname{ce}_m'(z, q) dz = 0, \quad (4.14в)$$

$$\int_0^{2\pi} \operatorname{se}_m(z, q) \operatorname{se}_m'(z, q) dz = 0, \quad (4.14г)$$

$$\int_0^{2\pi} \operatorname{ce}_m(z, q) \operatorname{se}_m'(z, q) dz = K_{mm} \approx m\pi, \quad (4.14д)$$

$$\int_0^{2\pi} \operatorname{se}_m(z, q) \operatorname{ce}_m'(z, q) dz = -K_{mm} \approx -m\pi, \quad (4.14е)$$

где приближенные выражения для величины K_{mm} в формулах (4.14д) и (4.14е) справедливы в интересующем нас интервале, получаем

$$(\alpha + \mu^2 - a_m)\lambda + 2\mu m = 0, \quad (4.15a)$$

$$-2\mu m\lambda + (\alpha + \mu^2 - b_m) = 0. \quad (4.15b)$$

В случае когда запрещенная зона достаточно узка и когда для энергии, лежащей внутри запрещенной зоны, выполняется неравенство $b_m \leq \alpha \leq a_m$, можно показать, что уравнения (4.15a) и (4.15b) аппроксимируются очень простым матричным уравнением

$$\begin{bmatrix} \alpha - a_m & 2\mu m \\ -2\mu m & \alpha - b_m \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = 0, \quad (4.16)$$

которым мы и будем здесь пользоваться вплоть до п. 5 § 4.

Секулярное уравнение, отвечающее этой матрице, дает соотношение для энергий

$$\alpha_{\pm} = \frac{1}{2} \{a_m + b_m \pm [(a_m - b_m)^2 - 16m^2\mu^2]^{1/2}\}. \quad (4.17)$$

Когда поверхность отсутствует, поверхностных состояний быть не должно, и величина μ исчезает; тогда эти энергии соответствуют просто краям зон $\alpha_+ = a_m$ и $\alpha_- = b_m$. Наличие поверхности, с другой стороны, допускает значение $\mu > 0$. Это аналогично поверхностному «возмущению», которое *стягивает* значения α_+ и α_- друг к другу и одновременно к центру запрещенной зоны. Такое притяжение отвечает знаку «минус» в подкоренном выражении в формуле (4.17). В свою очередь знак «минус» является следствием антисимметрии матрицы в уравнении (4.16). Выражение (4.17) в том виде, в котором оно выписано, излишне громоздко, в частности из-за наличия двух решений. Можно, обратив (4.17), получить

$$\mu = (2m)^{-1} [(a_m - \alpha)(\alpha - b_m)]^{1/2}. \quad (4.18)$$

К такой же формуле можно прийти непосредственно из уравнения (4.16). По определению для поверхностного состояния перед корнем должен быть взят только знак «плюс».

Из (4.16) можно также получить параметр гибридизации

$$\lambda = \frac{\alpha - b_m}{2\mu m}. \quad (4.19)$$

При этом величина λ должна быть положительной, поскольку параметр μ по определению положителен и $b_m < \alpha < a_m$. Это утверждение эквивалентно «правилу отбора для поверхностного состояния» $\lambda > 0$, важные следствия которого будут обсуждаться ниже. Объединяя выражение (4.18) и (4.19), получаем

$$\lambda = \left(\frac{\alpha - b_m}{a_m - \alpha} \right)^{1/2}. \quad (4.20)$$

Формулы (4.18) и (4.20) могут быть преобразованы так, чтобы они определяли величины α и μ как однозначные функции параметра λ :

$$\alpha = \frac{b_m + \lambda^2 a_m}{1 + \lambda^2}, \quad (4.21a)$$

$$\mu = \frac{\lambda (a_m - b_m)}{2m (1 + \lambda^2)}. \quad (4.21b)$$

Рассмотрим теперь следствия из соотношений (4.21a) и (4.21b), считая λ независимой переменной. Ниже параметр λ будет определен путем детального анализа условий на поверхности. Когда $\lambda = 0$, то $\alpha = b_m$ и $\mu = 0$. Энергия отвечает дну запрещенной зоны, и поверхностное состояние отсутствует. Если $\lambda = 1$, то $\alpha = (a_m + b_m)/2$, $\mu = (a_m - b_m)/4m$, и поверхностное состояние отвечает уровню энергий, расположенному в центре запрещенной зоны, причем декремент затухания его волновой функции (или степень локализации) максимален. Когда $\lambda = \infty$, то $\alpha = a_m$ и $\mu = 0$. Вблизи верхних и нижних краев зон энергия изменяется пропорционально квадрату волнового вектора μ , описывающего затухание, что следует из (4.17) и (4.18). Все эти результаты кажутся разумными и весьма удобными с аналитической точки зрения. Они справедливы при любых значениях параметров m и q , характеризующих объемные свойства. В п. 5 § 4 будет показано, что при учете

гибридизации с другими зонами эти результаты должны быть частично видоизменены.

Следует еще остановиться на случае, когда величина λ отрицательна. Из (4.19) и (4.21б) очевидно, что параметр μ также должен быть отрицателен. Этот случай (отрицательные λ и μ) в точности соответствует запрещенной волновой функции $\psi_i^*(z, q)$ из (4.12). Таким образом, отрицательным значениям λ отвечает физически нереальная ситуация, и эти значения должны быть исключены. Наконец, когда λ — чисто мнимая величина, энергия α лежит внутри запрещенных зон (4.20), и параметр μ — тоже чисто мнимый. При этом мы получаем блоховскую волну, которая не является поверхностным состоянием и здесь может не рассматриваться.

3. Условие непрерывности

Чтобы получить поверхностное состояние, необходимо выбрать такое значение энергии в объеме α , которому одновременно отвечает выполнение определенного условия непрерывности для волновой функции. Рассмотрим это условие несколько более подробно. Комбинируя соотношения (4.10) и (4.11), получаем общее условие непрерывности

$$\lambda = \frac{se_m(z_0, q) - (\xi + \mu)^{-1} se'_m(z_0, q)}{ce_m(z_0, q) - (\xi + \mu)^{-1} ce'_m(z_0, q)}. \quad (4.22)$$

Оно связывает параметр λ с индексом запрещенной зоны m , положением поверхности кристалла z_0 , амплитудой кристаллического потенциала $2q$ и высотой потенциальной ступеньки ξ , как показано на фиг. 20. Согласно результату п. 2 § 4, поверхностное состояние возможно только при условии, если параметр λ положителен. Следовательно, поверхностное состояние существует, если положительна правая часть уравнения (4.22). В общем случае связь величин m , z_0 , ξ , μ и q с параметром λ довольно сложная. Однако для некоторых частных примеров она может быть заметно упрощена даже без использования численного анализа.

В качестве первого примера рассмотрим случай, когда $z_0 = 0$, что соответствует обрыву кристалла в

точке *максимума* потенциала в элементарной ячейке. Полагая в (4.22) $z_0 = 0$, можно видеть, что величина λ всегда отрицательна [34]. Таким образом, для синусоидального потенциала, оканчивающегося в точке максимума, поверхностных состояний быть не может. В п. 7 § 4 будет установлена связь этого случая с состояниями шоклиевского типа [17].

Подобным же образом рассмотрим случай $z_0 = \pi/2$. Решетка при этом оканчивается в точке *минимума*. Подставляя в (4.22) $z_0 = \pi/2$, можно показать, что поверхностные состояния отсутствуют для четных m и будут существовать, когда m принимает нечетные значения. Как показано на фиг. 22, б, случаи « m равно четному числу» и « m равно нечетному числу» непосредственно отвечают центру и грани одномерной зоны Бриллюэна соответственно. Состояния будут находиться тем ближе к вершине запрещенной зоны, чем больше отношение ξ/m . Для бесконечной потенциальной ступеньки $\xi \rightarrow \infty$; таким образом, $\lambda \rightarrow \infty$, и поверхностное состояние исчезает, сливаясь с вершиной запрещенной зоны. О связи с таммовскими состояниями речь будет идти в п. 6 § 4.

При других положениях границы z_0 должно быть проведено более общее рассмотрение. Самый простой случай, когда $q \sim 0$ и $\xi \sim \infty$, отвечает почти пустой решетке и бесконечно высокому поверхностному потенциалу. При этом уравнение (4.22) сводится приближенно к виду

$$\lambda = \operatorname{tg} mz_0. \quad (4.23)$$

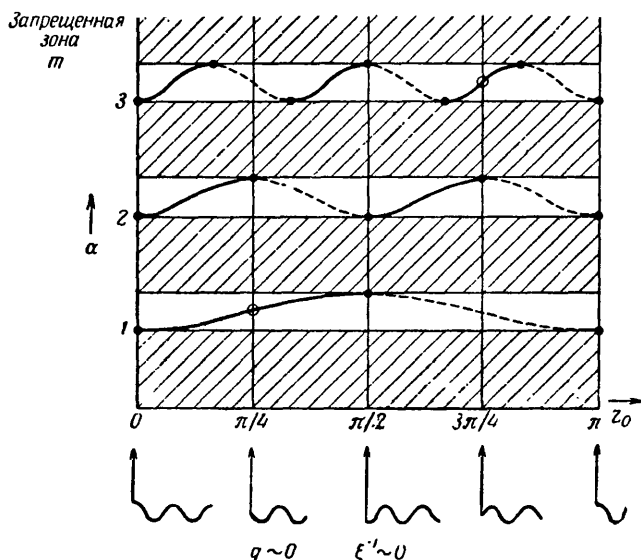
Правая часть (4.23) положительна только при

$$0 < mz_0 < \frac{\pi}{2}, \quad \pi < mz_0 < \frac{3\pi}{2}, \quad 2\pi < mz_0 < \frac{5\pi}{2}, \quad \dots \quad (4.24)$$

Таковы необходимые условия возникновения поверхностных состояний. Теперь с помощью соотношений (4.21а) и (4.23) может быть получено выражение для энергии:

$$\alpha = b_m \cos^2 mz_0 + a_m \sin^2 mz_0. \quad (4.25)$$

Эти результаты удобно представить графически, как это сделано на фиг. 23. Здесь построен график зависимости энергии α от координаты z_0 в интервале $0 < z_0 < \pi$. Поверхностные состояния в различных запрещенных зонах показаны синусоидальными сплошными кривыми.



Фиг. 23. Энергии поверхностных состояний как функции z_0 для почти пустой решетки и бесконечного поверхностного потенциала.

Рассмотрим, например, первую запрещенную зону ($m = 1$). Согласно (4.24), поверхностные состояния разрешены в интервале $0 < z_0 < \pi/2$. Одно из таких состояний отмечено кружком при $z_0 = \pi/4$. Пока координата z_0 смещается к значению $\pi/2$, энергия непрерывно нарастает от значения, отвечающего дну, до значения, соответствующего вершине запрещенной зоны. Однако для $z_0 > \pi/2$ параметр λ становится отрицательным, что соответствует *запрещенному* поверхностному состоянию; это отмечено пунктирной линией. Такое *запрещенное* состояние имеет правильное собственное значение α и

удовлетворяет условию непрерывности для волновой функции, однако оно *неправильно* ведет себя при $z \rightarrow \infty$. Тем не менее по соображениям симметрии нам будет удобно всегда включать эти запрещенные состояния, обозначая их на диаграммах пунктиром. Аналогичным образом вторая запрещенная зона ($m=2$) содержит поверхностные состояния в двух интервалах: для $0 < z_0 < \pi/4$ и $\pi/2 < z_0 < 3\pi/4$. Вообще m -я запрещенная зона имеет m различных областей поверхностных состояний. При $z_0 = 0$ все энергии α совпадают с нижними краями зон b_1, b_2, b_3, b_4 и т. д. При $z_0 = \pi/2$ энергии α чередуются: a_1, b_2, a_3, b_4 и т. д. Пожалуй, самым примечательным свойством фиг. 23 является ее сложная структура. Эту структуру было бы трудно объяснить, пользуясь простыми классическими соображениями; она возникает в основном из-за квантовомеханического условия непрерывности.

Более общий случай может быть исследован, если считать $q \sim 0$, а скачок ξ — конечным. Этот случай отвечает почти пустой решетке с конечным потенциальным барьером на поверхности. Уравнение (4.22) тогда принимает вид

$$\lambda = \frac{\sin mz_0 - [m/(\xi + \mu)] \cos mz_0}{\cos mz_0 + [m/(\xi + \mu)] \sin mz_0}. \quad (4.26)$$

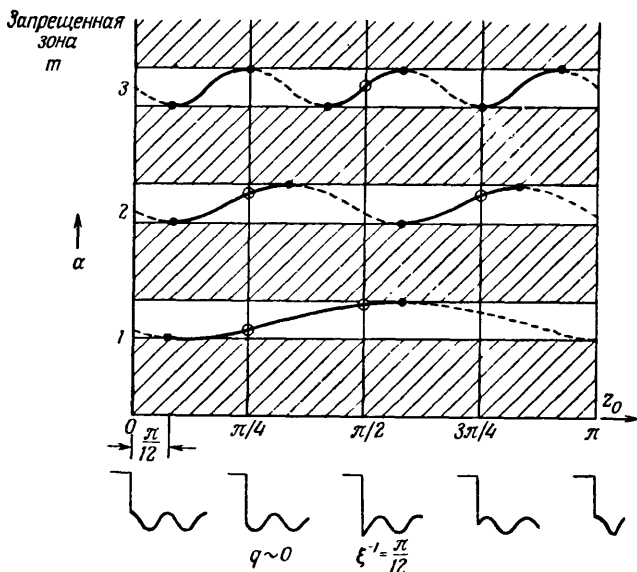
Когда $\xi \gg \mu$ и $m/\xi \ll 1$, что отвечает большому, но конечному барьеру, по отношению к предыдущим результатам новые члены могут рассматриваться как возмущение. Пользуясь тригонометрическим тождеством, можно свести (4.26) к выражению

$$\lambda = \operatorname{tg} m(z_0 - \xi^{-1}). \quad (4.27)$$

Единственная особенность этого уравнения по сравнению с (4.23) состоит в удалении слоя толщины ξ^{-1} с поверхности кристалла (появление фазового сдвига).

Поверхностные состояния, отвечающие (4.27), можно получить непосредственно однородным сдвигом фиг. 23 вправо на расстояние ξ^{-1} . Результат показан на фиг. 24, где для величины ξ^{-1} принято значение $\pi/12$. Более высокие зоны при $m > 2$ воспроизводятся менее точно, так как для них неравенство $m/\xi \ll 1$ не выпол-

няется. Важным результатом здесь является то, что для конечного барьера ξ поверхностные состояния появляются при $z_0 = \pi/2$, как это показано на диаграмме кружком. В соответствии с предыдущим расчетом для $z_0 = \pi/2$ такие состояния есть только в запрещенных



Фиг. 24. Энергии поверхностных состояний как функции z_0 для почти пустой решетки и конечного поверхностного потенциала.

Обратите внимание на сдвиг фазы (ξ^{-1}).

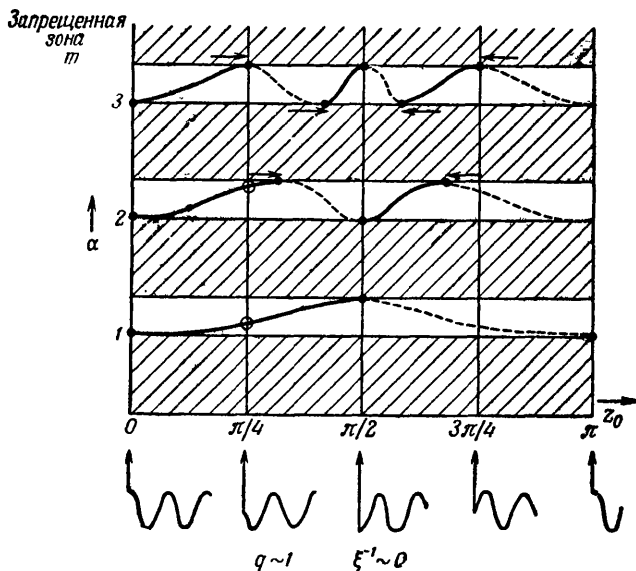
зонах с четными индексами m . Заметим также, что, как и в предыдущем случае, поверхностные состояния отсутствуют при $z_0 = 0$. Состояния, отвечающие значениям $m=2$, $z_0 = \pi/4$ и $m=2$, $z_0 = 3\pi/4$, будут называться «ионными поверхностными состояниями» по причине, указанной далее.

Свойства кристалла при $q > 0$ легко могут быть рассмотрены подобным же образом. Пусть сначала $q > 0$ и $\xi \rightarrow \infty$; это отвечает кристаллу с сильным периодическим потенциалом и бесконечным потенциалом у поверхности. Так как метод почти свободных электронов

теперь неприменим, воспользуемся непосредственно свойствами функций Матье. В этом случае вместо уравнения (4.22) будем иметь

$$\lambda = \frac{\text{se}_m(z_0, q)}{\text{ce}_m(z_0, q)}. \quad (4.28)$$

Как мы видели из фиг. 21, основной эффект увеличения параметра q состоит в «стягивании» максимумов и уз-



Фиг. 25. Энергии поверхностных состояний как функции z_0 для сильного решеточного и бесконечного поверхностного потенциалов.

Заметно сжатие в направлении точки $\pi/2$.

лов функций $\text{se}_m(z, q)$ и $\text{ce}_m(z, q)$ к точке $z = \pi/2$, где потенциал притяжения имеет наибольшее значение. Таким образом, координата границы z_0 , отвечающая значениям $\lambda = 0$ или ∞ , с ростом q будет сдвигаться ближе к точке $z_0 = \pi/2$. Иными словами, простым «сжатием» фиг. 23 можно получить фиг. 25. Сдвиг нулей на фиг. 25 условно показан стрелками. Здесь принято значение $q = 1$.

Наконец, рассмотрим особенно важный пример, когда в более общем случае $q > 0$ и ξ — конечная величина. Уравнение (4.22) можно теперь представить в другой форме, если использовать разложение в ряд без каких бы то ни было упрощений. Когда m принимает нечетные значения [34, 91], получаем

$$\lambda = \frac{\sum_{r=0}^{\infty} B_{2r+1}^{(2n+1)} [\sin(2r+1)z_0 - (2r+1)(\xi + \mu)^{-1} \cos(2r+1)z_0]}{\sum_{r=0}^{\infty} A_{2r+1}^{(2n+1)} [\cos(2r+1)z_0 + (2r+1)(\xi + \mu)^{-1} \sin(2r+1)z_0]} \quad (4.29)$$

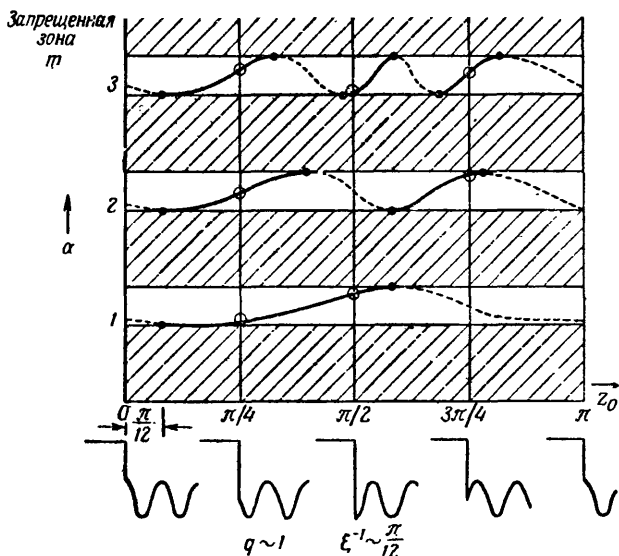
Если величина ξ достаточно велика, так что $\xi \gg \mu$ и $(2r+1)/\xi \ll 1$ для всех интересующих нас значений r (при условии, что $m/\xi \ll 1$), то в этом случае

$$\lambda = \frac{se_{2n+1}[(z_0 - \xi^{-1}), q]}{ce_{2n+1}[(z_0 - \xi^{-1}), q]} \quad (4.30)$$

Аналогичное соотношение можно получить и для четных значений m . Таким образом, данный случай *объединяет* эффекты сдвига вправо на величину ξ^{-1} (фиг. 23) и сжатия к центру на величину, пропорциональную q . Этот результат для $\xi = 12/\pi$ и $q = 1$ показан на фиг. 26. Необходимо отметить, что при $z_0 = 0$ поверхностные состояния опять отсутствуют, а при $z_0 = \pi/2$ они существуют, если m — нечетное число. Кроме того, для $m = 2$ и $z_0 = \pi/4$ ионное поверхностное состояние, впервые показанное на фиг. 24, смещено по направлению к *нижнему* краю зоны, а для $m = 2$ и $z_0 = 3\pi/4$ другое ионное поверхностное состояние расположено ближе к *верхнему* краю.

Резюмируя, можно сказать, что требование непрерывности налагает на параметры z_0 , q , ξ и m некоторые условия, необходимые для возникновения поверхностных состояний. Общая схема, представленная на фиг. 26, может быть получена с помощью свойств функций Матье как результат *сдвига* (ξ^{-1}) и *сжатия* ($\sim q$) более простой схемы фиг. 23. Физически наиболее важные поверхностные состояния отвечают значениям $z_0 = 0$,

$\pi/4$, $\pi/2$ и $3\pi/4$ и $m = 1, 2$. Связь этих состояний с поверхностными состояниями таммовского и шоклиев-



Фиг. 26. Энергии поверхностных состояний как функции z_0 для сильного решеточного и конечного поверхностного потенциалов. Ковалентным состояниям соответствуют нечетные значения m и $z_0 = \pi/2$, ионным состояниям — $m=2$ и $z_0 = \pi/4, 3\pi/4$.

ского типов, а также с ионными состояниями будет показана соответственно в п. 6, 7 и 8 § 4.

4. Конечный кристалл

Поверхностные состояния конечного кристалла должны обладать свойствами, несколько отличными от свойств полубесконечного кристалла, рассмотренного ранее. С физической точки зрения разумно ожидать, что полубесконечная модель будет весьма точным приближением для очень большого конечного кристалла ¹⁾. Мы

¹⁾ В работе [337] поверхностные и объемные состояния в тонкой пленке рассмотрены в рамках метода псевдопотенциала. — Прим. перев.

дадим здесь количественную оценку этих качественных эффектов. Можно показать [34], что в первом порядке теории возмущений сдвиг энергии $\Delta\alpha$, обусловленный конечностью решетки, содержащей N ячеек, в отличие от бесконечного кристалла определяется соотношением

$$\Delta\alpha \sim \omega \exp(-2\pi N\mu). \quad (4.31)$$

Таким образом, для достаточно больших значений N $\Delta\alpha \rightarrow 0$ и мы возвращаемся к предыдущим результатам. В случае тонких кристаллов при малых N , слабом затухании (при малых μ) и большом поверхностном барьере (при больших ω) сдвигом $\Delta\alpha$ больше нельзя пренебрегать и необходимо проводить дополнительные вычисления.

5. Гибридизация всех состояний

Рассмотрим наиболее общее поверхностное состояние $\psi_i(z, q)$, которое является суперпозицией всех объемных состояний типа стоячих волн [34, 91, 119], продолженных аналитически в запрещенную зону

$$\psi_i(z, q) = \sum_{r=0}^{\infty} [C_r \operatorname{se}_r(z, q) - S_r \operatorname{se}_r(z, q)] e^{-\mu z}. \quad (4.32)$$

Здесь коэффициенты C_r и S_r являются функциями q для каждой запрещенной зоны с номером r . Пусть поверхностное состояние расположено в m -й запрещенной зоне. Главными членами в сумме будут C_m и S_m , которые нами уже были рассмотрены ($\lambda = C_m/S_m$). Наша цель теперь — обсудить вклад состояний, отвечающих другим запрещенным зонам. При этом нет необходимости требовать, чтобы запрещенная зона была узкой; результаты в равной мере будут относиться ко *всем* зонам фиг. 22, а. Следуя процедуре, которая была использована при выводе (4.16), получим в нашем общем случае уравнение (4.33) [приведено на стр. 99].

Здесь введены обозначения

$$\alpha' = \alpha + \mu^2, \quad (4.34a)$$

$$H_{rr} = 2\pi\mu^{-1}K_{rr}, \quad (4.34б)$$

$$h_{rt} = 2\pi\mu^{-1}K_{rt}. \quad (4.34в)$$

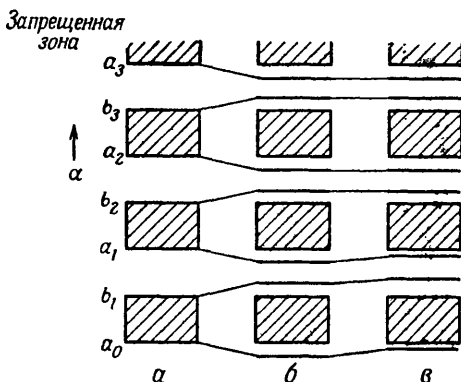
Можно показать, что $K_{rt} < K_{rr}$. В частности, когда $q \rightarrow 0$, $h_{rt} = 0$, но $H_{rr} \neq 0$.

Тот факт, что $h_{rt} = 0$, когда $r + t$ — нечетное число, имеет одно интересное физическое следствие. Поверхностное состояние, отвечающее точке Γ зоны Бриллюэна (четные запрещенные зоны, $m = 2, 4, 6, \dots$), представляет собой линейную комбинацию невозмущенных стоячих волн, каждая из которых относится к точке Γ ; невозмущенным стоячим волнам в точке X (нечетные запрещенные зоны, $m = 1, 3, 5, \dots$) отвечают исчезающие матричные элементы. Точно так же поверхностное состояние в точке X зоны Бриллюэна построено из невозмущенных стоячих волн для точки X , а невозмущенные состояния, отвечающие Γ , не дают вклада. Как можно видеть, элементы субматриц в (4.33), обведенных пунктирными линиями и расположенных вдоль диагонали, дают основной вклад. Проведенный ранее анализ эквивалентен пренебрежению малыми перекрестными членами h_{rt} . Детерминант (4.33) сводится в этом случае просто к произведению 2×2 -детерминантов, по одному для каждого m , за исключением $m = 0$. Пренебрежение матричными элементами h_{rt} ведет поэтому к разделению всех запрещенных зон, и проблема поверхностных состояний для m -й запрещенной зоны решается независимо от других. Кроме того, каждый из 2×2 -детерминантов будет иметь вид (4.16) в пределе узкой запрещенной зоны, когда $\alpha' \sim \alpha$, т. е. когда величиной μ^2 в диагональных матричных элементах (4.33) можно пренебречь. Учет члена μ^2 означает учет эффекта параболичности зонной структуры. При этом все эффективные массы электронов становятся меньше, чем все эффективные массы дырок. Исключение составляет низший уровень, для которого член μ^2 должен быть учтен уже в нулевом приближении¹⁾.

Главным следствием антисимметрии недиагональных членов H_{rr} , как было показано, является *притяжение*

¹⁾ Утверждение Левина [34], что уровень a_0 превращается в металлическое поверхностное состояние внутри первой разрешенной зоны, ошибочно. В этой работе при вычислении коэффициента a_0 был опущен член μ^2 в диагональном элементе (4.33).

краев запрещенных зон друг к другу. Это не относится к основному уровню a_0 , который ввиду наличия члена μ^2 в коэффициенте α' понижается. Схематически это показано на фиг. 27, *а* и *б*. Следует отметить, что энергетические уровни, показанные на фиг. 27, — это только *возможные* уровни, получаемые с помощью аналитического продолжения. Чтобы им отвечали истинные поверхностные состояния, соответствующие волновые



Фиг. 27. Схема энергетических уровней.

α — объемные состояния; β — возможные поверхностные состояния в приближении узкой зоны; γ — возможные поверхностные состояния при учете межзонных членов притяжения.

функции должны иметь положительные значения μ и удовлетворять условию непрерывности в точке $z = z_0$.

Перекрестные члены h_{rt} нарушают независимость запрещенных зон. Не проводя детальных расчетов, можно оценить качественно результат учета элементов h_{rt} , исходя из того, что матричные элементы антисимметричны, поэтому перекрестные члены стремятся *стянуть* вместе уровни энергии. Частный случай подобного рода был подробно рассмотрен на основе уравнения (4.16) для изолированной запрещенной зоны. Сдвиг уровней под действием притяжения показан на фиг. 27, *в*.

Больше всех подвержен возмущению подобного типа уровень a_0 , отвечающий дну первой разрешенной зоны. Перекрестный матричный элемент h_{02} вызовет сдвиг

этого уровня в сторону больших энергий (фиг. 27, в), тогда как элемент, который мог бы оказать противоположное действие, отсутствует. Подобным же образом ведут себя и гибриды состояний a_1 и b_1 . Они притягиваются в результате присутствия членов h_{13} и h_{31} , и их энергии несколько повышаются, как это показано на фиг. 27, в. Вторая и последующие зоны, вероятно, не возмущаются сколько-нибудь сильно, так как притяжение действует в обоих противоположных направлениях.

Таким образом, при использовании полной матрицы (4.33) вместо 2×2 -матрицы (4.16) были получены новые результаты, однако число их невелико. Поэтому схемы на фиг. 23—26 существенно не меняются.

6. Таммовские состояния

Таммовские состояния [14] рассматриваются большинством авторов как особый тип поверхностных состояний. Их появление обычно связывают с сильным поверхностным «возмущением», а исчезновение — с отсутствием такого возмущения. Это определение представляется туманным и иногда приводит к заблуждениям, поскольку само наличие поверхности всегда означает довольно значительное возмущение.

Аналогом *сильного поверхностного возмущения* одноатомного кристалла в задаче Маттье должен быть обрыв кристалла, скажем, в точке *минимума* потенциала, что отвечает значению $z_0 = \pi/2$ на фиг. 20 и 26. Предполагается, что атомы в одноатомном кристалле находятся в местах минимума потенциала: $z = \pi/2, 3\pi/2, 5\pi/2, \dots$. Как показано на фиг. 26, при $z_0 = \pi/2$ следует ожидать появления таммовских поверхностных состояний только в нечетных запрещенных зонах. Поверхностное возмущение оказывается сильным только для некоторых запрещенных зон, что согласуется со схемой Тамма. Другими словами, для синусоидального потенциала таммовские состояния появляются только в точке X одномерной зоны Бриллюэна, как показано на фиг. 22, б. Вывод о том, что в задаче Маттье таммовские состояния появляются в нечетных запрещенных зонах, впервые был сделан Статцем [21], который исходил из

аналогии между диаграммой пересекающихся зон Шокли [17] и схемой зависимости энергии α от параметра q (фиг. 22, а). Статц не рассматривал, однако, условия непрерывности и не получал для энергетических уровней выражений типа (4.21а) или (4.22). Появление таммовских состояний в запрещенной зоне с $m = 1$ также можно предсказать на основе метода почти свободных электронов, следуя Мауэ [15] и Гудвину [16]. Согласно этому подходу, поверхностное состояние возникает в m -й запрещенной зоне, если m -я фурье-компонента потенциала отрицательна (m -я гармоника обрывается в точке минимума). Переводя это на язык таммовских состояний в задаче Матье, имеем $z_0 = \pi/2$, $V_1 < 0$, $V_2 = V_3 = V_4 = \dots = 0$, и появление таких состояний следует ожидать при $z_0 = \pi/2$ в первой запрещенной зоне с $m = 1$. Это состояние показано на фиг. 26. Подход почти свободных электронов ничего не говорит о проблеме таммовских поверхностных состояний в запрещенной зоне с $m > 1$, поскольку гармоники, высшие по отношению к синусоиде, фигурирующей в задаче Матье, отсутствуют. Иначе говоря, зависимость энергии α от параметра q на фиг. 22, а оказывается линейной при $m = 1$, однако если $m > 1$, то α изменяется с q по квадратичному закону. Линейная зависимость может быть описана теорией почти свободных электронов, тогда как квадратичная требует привлечения теории возмущений с учетом членов более высокого порядка. В модели прямоугольных ям, которые обладают значительным числом высших гармоник, поверхностные состояния рассмотрены Статцем [21].

В задаче Матье аналогом случая, когда *отсутствует поверхностное возмущение*, в одноатомном кристалле служит потенциал, оборванный в точке максимума $z_0 = 0$. (Ступенька на поверхности оставляет, строго говоря, неисчезающее возмущение.) Таким образом, таммовские состояния должны *отсутствовать* для поверхности при $z_0 = 0$, как показано на фиг. 26. Этот результат, предсказанный Статцем [21], можно также ожидать в случае $m = 1$ для метода почти свободных электронов, который использовал Гудвин. Ранее были приведены гармоники потенциала, вычисленные для

$z_0 = 0$: $V_1 > 0$, $V_2 = V_3 = V_4 = \dots = 0$. Так как $V_1 > 0$, в первой запрещенной зоне при $z_0 = 0$ поверхностных состояний не возникает.

В свете всего вышесказанного, вероятно, будет полезно сравнить результаты, полученные в задаче Маттье, с результатами методов сильной связи и кристаллических орбиталей, обсуждавшимися в гл. 3. Для этого нам необходимо согласовать терминологию. В одном случае поверхность характеризуется положением обрыва внутри последней элементарной ячейки, z_0 , фазовым сдвигом ξ^{-1} , связанным с высотой ступеньки на поверхности, и константой затухания волновой функции внутри кристалла μ . В представлении ЛКАО поверхность описывается возмущенным кулоновским интегралом α' , отличающимся от кулоновского интеграла в объеме α , и константой затухания μ (ЛКАО). Сопоставить теперь терминологию можно, приравняв коэффициенты μ . В случае поверхностных состояний вблизи дна запрещенной зоны ($\lambda \sim 0$) в задаче Маттье имеем

$$\mu = \frac{1}{2} (a_m - b_m) (z_0 - \xi^{-1}). \quad (4.35)$$

В то же время для задачи ЛКАО с одной зоной имеем, согласно Гудвину [16],

$$\mu \text{ (ЛКАО)} = \pi^{-1} [\ln (\alpha' - \alpha) - \ln \beta]. \quad (4.36)$$

Таким образом, получается соответствие [34] между этими двумя подходами:

$$\alpha - \alpha' = \exp \left[\frac{\pi z_0 (a_m - b_m)}{2} \right], \quad (4.37)$$

$$\beta = \exp \left[\frac{\pi (a_m - b_m)}{2\xi} \right]. \quad (4.38)$$

Физически разность кулоновских энергий $\alpha' - \alpha$ в терминологии ЛКАО связана с положением точки обрыва z_0 , как это видно из соотношения (4.37). Положительная величина $\alpha' - \alpha$ соответствует увеличению поверхностного потенциала отталкивания в последней ячейке. Это оказывает такое же действие, что и сдвиг координаты z_0 в положительном направлении, как изображено на фиг. 26. Аналогично критическое значение энергии

при возникновении поверхностных состояний ($\alpha' - \alpha \geq \beta$) связано с высотой ступеньки ξ на поверхности (или с фазовым сдвигом) согласно (4.38). Эти результаты справедливы для всех m . Почти такое же рассмотрение можно привести в случае поверхностных состояний, находящихся у вершины запрещенной зоны ($\lambda \rightarrow \infty$).

7. Шоклиевские состояния

Шоклиевские состояния [17] рассматриваются большинством авторов как другой частный тип поверхностных состояний, альтернативный по отношению к таммовским состояниям. Обычно считается, что шоклиевские состояния возникают в отсутствие поверхностного возмущения и обусловлены пересечением зон. Эти состояния исчезают с появлением сильного поверхностного возмущения. Как отмечалось ранее, такое определение туманно и несколько двусмысленно. В одномерной задаче с периодическим потенциалом сдвиг на половину ячейки (т. е. новое определение координат атомных ядер) иногда превращает таммовские состояния в шоклиевские и наоборот. Это обстоятельство было со всей ясностью подчеркнуто Липпманом [24].

Аналогом слабого поверхностного возмущения в задаче Матье для одноатомного кристалла должен быть, например, потенциал, оборванный в точке *максимума*, что отвечает значению $z_0 = 0$ на фиг. 20 и 26. Каким образом можно узнать, что в задаче Матье зоны пересекаются? Из фиг. 21 видно, что функции $se_m(z, q)$, $se_m(z, q)$ с ростом q непрерывно переходят в атомные функции. Эти функции обычным путем упорядочены по энергиям: связывающая s , антисвязывающая s , связывающая p , антисвязывающая p , связывающая d , ... и т. д. Поэтому можно думать, что на диаграмме зависимости α от q (фиг. 22, а) пересечение зон отсутствует, и, следовательно, шоклиевские состояния не возникают. Как показывает фиг. 26, шоклиевских состояний нет при $z_0 = 0$.

Вместе с тем нет, однако, никакой гарантии, что обрыв в точке максимума потенциала является наиболее подходящей моделью для одномерной ковалентной

решетки (такой, как в случае Si, Ge, алмаза), которая молчаливо принималась Шокли. Во-первых, вычисленный Клейнманом и Филлипсом [121] псевдопотенциал валентного электрона имеет *относительный минимум между ближайшими соседними атомами ковалентной решетки*. Этот минимум обуславливает концентрацию электронной плотности в промежутке *между* ковалентными атомами, что можно было предположить на основании простых химических соображений. Таким образом, более подходящим местом для обрыва ненасыщенной ковалентной связи может оказаться минимум потенциала [5]¹⁾. В задаче Матье это должно соответствовать обрыву при $z_0 = \pi/2$. В этом случае действительно есть поверхностные состояния, но они уже были ранее отнесены к таммовскому типу. Это усугубляет определенное неудобство при делении состояний на таммовские и шоклиевские. Во-вторых, как известно в настоящее время из экспериментов по ДЭНЭ, ковалентные кристаллы испытывают перестройку поверхности (она будет рассмотрена в гл. 7). Только что образованные ненасыщенные связи, по-видимому, настолько нестабильны, что поверхностные атомы спонтанно перестраиваются с образованием более стабильной конфигурации. В этом случае нельзя говорить о ненасыщенных связях в обычном смысле, так как они до некоторой (неизвестной) степени насыщены в результате атомной перестройки. Наконец, преобразование $q \rightarrow -q$ меняет порядок объемных волновых функций с $se_0, se_1, se_1, se_2, se_2, se_3, \dots$ на $se_0, se_1, se_1, se_2, se_2, se_3, \dots$ [91, 119] без изменения их энергий. Это выглядит как преобразование, одновременно сдвигающее начало координат на $\pi/2$ (добавление половины ячейки к кристаллу) и меняющее чередование зон. Таким образом, существует много поводов для недоразумений при определении шоклиевских состояний. Ситуация, по-видимому, несколько упрощается, если применить методы кристаллических орбиталей. В этом случае считается, что состояниям типа Шокли отвечают невозмущенные кулоновский и

¹⁾ J. D. Levine, не опубликовано. См. также [122].

резонансный интегралы на поверхности и пересекающиеся зоны [2, 16, 22, 123].

Поскольку классификация поверхностных состояний оказывается зависящей от типа рассматриваемого вещества (ионные, ковалентные, металлические кристаллы и т. д.), вероятно, будет благоразумнее отказаться от создания какой-либо общей схемы, подобной делению на таммовские и шоклиевские состояния, и рассматривать каждый из этих типов веществ по отдельности. При этом можно избежать двусмысленности, простирающейся из несколько неудачного определения. В п. 8 § 4 дан пример одной из таких возможных схем.

8. Ионные поверхностные состояния

Ионные и частично ионные кристаллы, такие, как соединения элементов групп III—V, II—VI и I—VII, обладают сходным расположением собственных поверхностных состояний (п. 3 § 17). Эти кристаллы относятся к типу АВ, где А — металлический катион, а В — анион неметалла. Обозначение АВ полезно даже в случае кристаллов, имеющих, вообще говоря, дробный заряд. Детальное рассмотрение ионных поверхностных состояний в п. 2 § 1, а также в § 10 и в п. 3 § 17 показывает, что объемная зона проводимости (валентная зона) может рассматриваться как образованная преимущественно из состояний ионов типа А(В). Ширина запрещенной зоны определяется электростатической, или маделунговской, энергией. Левин и Марк [30] рассчитали маделунговскую энергию для поверхностных ионов и с ее помощью получили «парное» расположение поверхностных состояний. Поверхностные состояния типа А(В) оказались расположенными на расстоянии в несколько десятых электрон-вольта ниже (выше) зоны проводимости (валентной зоны), если поверхностная плоскость содержит ионы обоих типов. Такие парно расположенные поверхностные состояния, показанные схематически в § 12 на фиг. 34, д, удобно назвать «ионными поверхностными состояниями», как мы и примем в дальнейшем. Можно надеяться, что эта терминология в случае образцов с чисто А- или В-поверхностями будет весьма

наглядной и значительно более полезной, чем разделение состояний на состояния таммовского и шоклиевского типа.

Ионные поверхностные состояния, возникающие и в задаче Матье, обладают аналогичными свойствами. Эти состояния могут быть исследованы с помощью фиг. 26, если учесть два следующих обстоятельства. Во-первых, модель кристалла типа АВ может быть рассмотрена в рамках задачи Матье, если принять, что ионы А(В) помещены в точках максимума (минимума) синусоидального потенциала. Такое соответствие означает, что ионы А(В) электроположительны (электроотрицательны) в согласии с обычными химическими представлениями. Обрыву кристалла на ионе А(В) теперь отвечает значение $z_0 = 3\pi/4$ ($z_0 = \pi/4$), что видно из фиг. 26. Этим точкам не соответствуют максимумы или минимумы потенциала, они просто делят пополам расстояние между ионами. Во-вторых, запрещенной зоне, разделяющей валентную зону и зону проводимости, для большинства соединений элементов III—V и II—VI групп отвечает точка Г ($k = 0$) зоны Бриллюэна. В задаче Матье это *вторая* запрещенная зона ($m = 2$), поскольку первая (точка X одномерной зоны Бриллюэна) разделяет *s*- и *p*-зоны иона В, заполненные согласно простой химической картине. Поверхностные состояния типа А(В) действительно существуют для $m = 2$ при $z_0 = 3\pi/4$ ($z_0 = \pi/4$), как это видно из фиг. 26. Можно, кроме того, показать, что с увеличением степени ионности (с ростом q) состояния А(В) стремятся к «своей» зоне проводимости (валентной зоне).

Очень полезно сравнить свойства ионных поверхностных состояний, полученных выше, с теми же свойствами в модели края зоны метода ЛКАО (§ 1)¹⁾. В последнем случае ионному кристаллу отвечает система *s(p)*-орбиталей, связанных с А(В) ионами. В трехмерном случае придется взять три ортогональные *p*-орбитали. Легче всего сопоставить оба метода, срав-

¹⁾ В работе [340] методом функций Грина исследованы поверхностные состояния у поверхности (110) CsBr. — *Прим. перев.*

нивая матрицы энергий, переписанные в обозначениях задачи Матье [34]

$$\begin{bmatrix} a_m - \alpha & -2\pi^{-1}K_{mm}\mu \\ 2\pi^{-1}K_{mm}\mu & b_m - \alpha \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = 0 \quad (4.39)$$

и в обозначениях модели края зоны метода ЛКАО [33]

$$\begin{bmatrix} \alpha_A - E & -2\beta \operatorname{sh} \mu \\ 2\beta \operatorname{sh} \mu & \alpha_B - E \end{bmatrix} \begin{bmatrix} R \\ 1 \end{bmatrix} = 0. \quad (4.40)$$

В уравнении (4.39) m — номер запрещенной зоны, равный 2 в случае ионных поверхностных состояний. Поскольку $a_2 = \alpha_A$ и $b_2 = \alpha_B$, диагональные элементы в обоих уравнениях одинаковы. Недиagonальные элементы в (4.39) содержат безразмерный интеграл $\mathbf{k} \cdot \mathbf{p}$ -типа, а именно:

$$K_{mm} = \int_0^{2\pi} \operatorname{se}_m(z, q) \frac{d}{dz} \operatorname{se}_m(z, q) dz. \quad (4.41)$$

Этот интеграл объединяет функции $\operatorname{se}_m(z, q)$ и $\operatorname{se}_m(z, q)$, которые служат волновыми функциями состояний, отвечающих соответственно вершине и дну m -й запрещенной зоны. Аналогично, недиагональные элементы в (4.40) содержат интеграл резонансного типа величины β . Он тоже связывает края зоны, но несколько иным образом. В любом случае недиагональный член β дает кривизну краев зоны, как показано на фиг. 2; таким образом, он может быть определен из экспериментов с эффективной массой.

В задаче Матье параметр λ описывает гибридизацию функций $\operatorname{se}_m(z, q)$ и $\operatorname{se}_m(z, q)$, тогда как в методе ЛКАО гибридизация ψ_A и ψ_B определяется параметром R .

Аналогия также прослеживается и в уравнениях сшивания на поверхности. Как λ , так и R однозначно определяются этими уравнениями, и условия существования поверхностных состояний принимают вид $\lambda > 0$ и $R > 0$ [33]. Эту аналогию можно продолжить даже в том случае, когда в задаче Матье требуется сшивание функции ψ и ее производной на поверхности в точке z_0 , тогда как в методе ЛКАО мы просто отбрасываем

атомы, находящиеся слева от поверхности ($a_0 = 0$, $a_{-1} = 0$, $a_{-2} = 0$ и т. д.).

Полезно также сравнить метод края зоны ЛКАО и задачу Маделунга [30]. В первом случае в качестве энергий краев зон берут просто значения α_A и α_B , а энергии поверхностных состояний α'_A и α'_B . Простота этих выражений является следствием недостаточно строгого с точки зрения квантовой механики рассмотрения задачи Маделунга. Классическая задача Маделунга тоже не дает уширения объемных зон, что вызвано отсутствием недиагональных членов, содержащих β .

Глава 5

РАСЧЕТЫ ТРЕХМЕРНЫХ СЛУЧАЕВ

До сих пор мы рассматривали только одномерные модели, причем такой выбор определялся их сравнительной простотой¹⁾. В этой главе нам предстоит изучить более реалистические случаи, учитывающие наличие атомных плоскостей, параллельных поверхности. При этом возникают трудности трех видов: во-первых, как обнаружено экспериментально, некоторые поверхности испытывают перестройку или хотя бы слабое перераспределение атомов; во-вторых, поведение кристаллического потенциала вблизи поверхности недостаточно известно и, наконец, в-третьих, всякий трехмерный расчет сопряжен с вычислительными трудностями.

Поверхностные атомы часто смещены в продольном и поперечном по отношению к поверхности направлениях (п. 3 § 1 и § 15). Поскольку для точного расчета необходимо знать положения атомов, то для того, чтобы учесть эти геометрические эффекты, приходится обращаться к экспериментальным данным. Если потенциал известен точно, то уравнение Шредингера для трехмерного конечного кристалла может быть в принципе решено с достаточной точностью без искусственного введения границы или условий сшивания, что, естественно, приводит к объемным и поверхностным состояниям. Практически, однако, приходится прибегать к идеализации строения приповерхностной области. Упрощающими предположениями здесь служат: скачок потенциала в виде ступеньки в одномерном и разрыв непрерывности на плоскости в трехмерном случаях [124]. С вычислительной точки зрения трехмерные расчеты объемной зонной структуры являются в настоящее время довольно привычной процедурой [98, 106]. Однако при наличии

¹⁾ В прямом методе кристаллических орбиталей трехмерная система сводится к одномерной при помощи так называемых «слоистых орбиталей», которые обсуждаются в п. 2а § 2.

поверхности возникают дополнительные сложности, связанные с граничными условиями, описывающими поверхностные эффекты, что делает задачу на порядок более трудной.

§ 5. Непосредственное обобщение одномерной задачи

К задаче о реальном трехмерном кристалле проще всего подойти, взяв одномерную модель и вводя затем два других измерения, если это возможно, путем некоторой «суперпозиции». Например, в методе кристаллических орбиталей (гл. 3) выражение для трехмерной приведенной объемной энергии идеализированного одноатомного кубического кристалла имеет вид

$$X = \cos \theta_x + \cos \theta_y + \cos \theta_z, \quad (5.1)$$

где θ_z — безразмерный волновой вектор, нормальный к поверхности, а $\theta_{x,y}$ — волновые векторы, параллельные ей. Соответствующая волновая функция запишется в следующей форме:

$$\psi = \psi(\theta_x) \psi(\theta_y) \psi(\theta_z). \quad (5.2)$$

Принимая во внимание наличие поверхности, положим $\theta_z = i\mu_z$ или $(\pi + i\mu_z)$, тогда вместо (5.1) имеем (1.66)

$$X = \cos \theta_x + \cos \theta_y \pm \operatorname{ch} \mu_z, \quad (5.3)$$

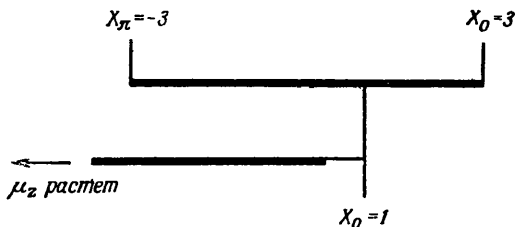
и волновая функция (5.2) теперь экспоненциально затухает в направлении z . Если $\theta_{x,y,z}$ или π , то $X_0 = 3$ и $X_\pi = -3$. Ширина объемной зоны поэтому имеет вид

$$\Delta X_v = X_0 - X_\pi = 6.$$

В формуле (5.3) минимальное значение μ_z , допускаемое условиями существования поверхностных состояний, есть $\mu_z = 0$ [33], что отвечает краю зоны. Взяв при этом знак «минус» в (5.3) и полагая $\theta_{x,y} = 0$ и π , так что $X_0 = -3$, получим следующее значение для ширины зоны поверхностных состояний:

$$\Delta X_s = 4.$$

Таким образом, зоны объемных и поверхностных состояний полностью перекрываются. С ростом μ_z поверхностная зона сдвигается влево (фиг. 28) и начинает отделяться от объемной. Хотя величина этого сдвига зависит от величины μ_z , сами ширины зон от нее не зависят: $X_v = 6$ и $X_s = 4$. Следует отметить, что в более общей



Фиг. 28. Перекрывание объемной зоны ($\Delta X_v = 6$) и поверхностной зоны ($\Delta X_s = 4$).

постановке трехмерной задачи величина μ_z сама оказывается функцией θ_x, y . Дальнейшее обсуждение перекрывания объемных и поверхностных зон и его связи с резонансным рассеянием будет дано в § 8.

§ 6. Аналитическое продолжение

Описанная выше простая процедура в принципе может быть также применена и в более сложных случаях. Как показали Хейне [124, 126], а также Бадро и Хейне [127], Блаут [128], Джонс [129], Стерн и др. [130], экспоненциально затухающие поверхностные состояния могут быть получены *аналитическим продолжением* объемной зонной структуры. При этом объемные энергии $E_n(\theta_x, \theta_y, \theta_z)$ и волновые функции $\Psi_n(\theta_x, \theta_y, \theta_z)$, отвечающие n -й зоне, должны быть вычислены для комплексных значений θ_z . Остается, однако, проблема граничных условий сшивания, и поэтому необходимо брать линейную комбинацию нескольких аналитически продолженных объемных волновых функций, часто дополненную другими, более локализованными волновыми

функциями¹⁾. Следует, конечно, иметь в виду, что аналитическое продолжение не зависит от способа описания волновой функции, будь то присоединенные плоские волны [131], $k \cdot r$ [132], метод сильной связи [2, 133], метод псевдопотенциала [134] или другие²⁾.

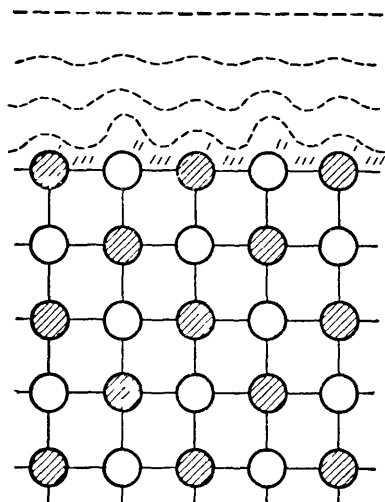
§ 7. Структура приповерхностной области

Даже самая лучшая схема зонной структуры, продолженная аналитически надлежащим образом, не дает гарантии, что полученные поверхностные состояния достаточно реалистичны. Это вызвано неопределенностью приближенного описания приповерхностной области. Для большинства методов (присоединенных плоских волн, псевдопотенциала, $k \cdot r$ -приближения и почти свободных электронов [16]) поверхность представляется плоским разрывом непрерывности. Как можно видеть с помощью фиг. 29, этим вносятся три основных упрощения: а) Геометрическое место точек равной энергии вблизи поверхности не является плоскостью, как это обычно предполагается, но отражает «зернистую» атомную структуру. Она показана пунктирными линиями [5]. б) Как показано на фигуре, каждому значению энергии отвечает свой собственный профиль эквиэнергетической поверхности. Точный расчет должен содержать сшивание волновых функций внутри и вне кристалла, соответствующих данной энергии вдоль всей волнообразной

¹⁾ Имеются в виду волновые функции $\Psi(\theta_x, \theta_y, \theta_z)$, у которых параллельная поверхности часть волнового вектора $(\theta_x^2 + \theta_y^2)^{1/2}$ выходит за пределы эквиэнергетической поверхности $X(\theta_x, \theta_y, \theta_z)$, так что перпендикулярная компонента θ_z автоматически оказывается комплексной. См. также работу [327]. — *Прим. перев.*

²⁾ В последние годы появилось несколько работ, использующих аналитическое продолжение для рассмотрения поверхностных состояний в металлах [319, 326—328, 334, 339]. В частности, было показано, что поверхностные состояния могут возникать не только в запрещенных зонах, отвечающих краям зоны Бриллюэна и обусловленных брэгговскими отражениями, но и в запрещенных зонах внутри зоны Бриллюэна, связанных, например, с $s-d$ -гибридизацией в переходных металлах. В работе [329] был рассчитан эффект туннелирования электронов из поверхностных состояний такого типа. Метод аналитического продолжения в рамках $k \cdot r$ -теории возмущений применялся в работе [330]. — *Прим. перев.*

поверхности постоянной энергии. в) Обычно используемое приближенное описание эквиэнергетического профиля плоским разрывом, не учитывающее описанной структуры [16, 131, 132, 134], приводит к сшиванию волновых функций лишь в нескольких симметричных точках поверхностной элементарной ячейки. Кроме того, само положение такой плоскости разрыва неизвестно.



Ф и г. 29. Схема профилей потенциальной энергии у поверхности двухатомной решетки.

Пунктирные линии отвечают геометрическому месту точек равной потенциальной энергии. Штрихами обозначены разные возмущенные кулоновские и резонансные интегралы метода кристаллических орбиталей в приближении сильной связи.

Роль этих упрощений можно, вероятно, свести к минимуму, прибегнув к методу кристаллических орбиталей в приближении сильной связи (гл. 3). В этом случае поверхностная область характеризуется несколькими *возмущенными* кулоновскими и резонансными интегралами. Последние могут иногда быть оценены с помощью простого подхода ненасыщенных связей (§ 9), методов классической электростатики (§ 10) или квантовой механики.

Между тем методы кристаллических орбиталей, по-видимому, не являются наилучшими для расчетов

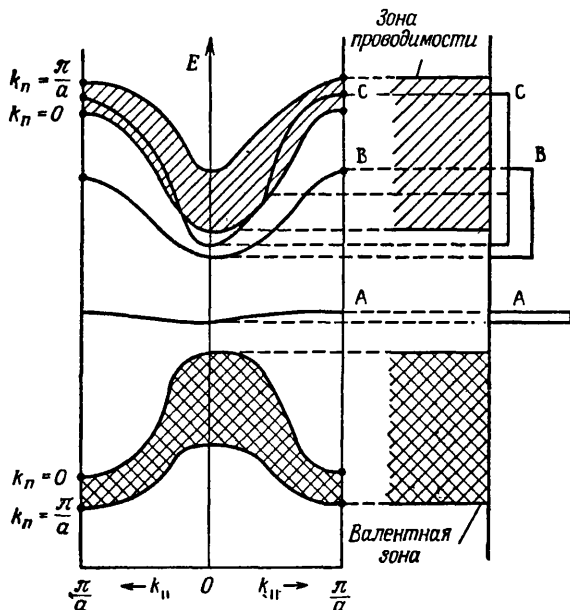
объемной зонной структуры. Иными словами, использование нескольких атомных орбиталей не позволяет адекватно описать зависимость E от k в энергетическом интервале, скажем, равном 20 эВ. Однако если мы интересуемся только поверхностными состояниями, принадлежащими одной запрещенной зоне шириной, например, 2 эВ, то применение методов кристаллических орбиталей более оправданно. Известные диаграммы краев зон могут быть аппроксимированы с помощью метода края зоны [33, 135]. Основная идея заключается в том, чтобы приписать известным краям зон определенные значения кулоновских интегралов, а известным значениям их кривизны — резонансные интегралы (эффективные массы электронов и дырок). Этот частный метод был с успехом применен в случае соединений элементов III—V и II—VI групп [135], в которых запрещенная зона между валентной зоной и зоной проводимости — прямая; ей отвечает точка Γ — центр зоны Бриллюэна [120]. Недавно путем сопоставления результатов, полученных этим методом, с экспериментальными данными Левин и Фриман [135] определили положения поверхностных атомов типа А и В ¹⁾.

§ 8. Плотность состояний

В случае одномерных расчетов из рассмотрения выпадает энергетическое распределение собственных поверхностных состояний, или *плотность состояний*. Число поверхностных состояний может оказаться порядка единицы на поверхностный атом. В трехмерном случае поверхностные электроны взаимодействуют между собой, и как следствие происходит уширение поверхностной зоны. Как энергия, так и плотность состояний могут быть в принципе наблюдаемы экспериментально (гл. 7). На фиг. 30 схематически показаны некоторые возможные простые примеры структуры зон поверхностных состояний и их распределений по энергии. Энергии объемных и поверхностных состояний построены как функции вол-

¹⁾ Оригинальный подход к рассмотрению структуры приповерхностной области с использованием правила сумм Фриделя применен в работе [335]. — *Прим. перев.*

нового вектора, параллельного плоскости поверхности. Объемные состояния, кроме того, характеризуются нормальным волновым вектором k_n , в результате чего объемная зона является квазиконтинуумом с запрещенной зоной в центре, как это имеет место для большинства соединений типа $A^{III}B^V$ и $A^{II}B^{VI}$ [120].



Фиг. 30. Энергия как функция волнового вектора и плотность состояний для некоторых возможных моделей поверхностных состояний.

В случае А (фиг. 30) большое число ($\sim 4 \cdot 10^{14}$ см $^{-2}$) поверхностных уровней сосредоточено в очень узком энергетическом интервале ($\sim 10^{-1}$ эВ). Отсюда плотность состояний составляет $\sim 4 \cdot 10^{15}$ см $^{-2}$ ·эВ $^{-1}$; это значение характерно для Ge и Si (гл. 7). В случаях В и С поверхностные зоны слегка смещены относительно краев объемных зон, и, как показано, объемные и поверхностные зоны могут в значительной степени перекрываться.

Отличие случая C от B состоит в том, что в первом происходит пересечение края объемной зоны, и это приводит к возникновению поверхностных резонансных состояний [135] вместо обычных локализованных. Кривая B похожа на кривую C тем, что ей отвечает приблизительно такая же кривизна в точке $k_{\parallel} = 0$ (т. е. близкая эффективная масса). Плотность состояний для относительной эффективной массы, равной 0,2, составляет $\sim 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$ (гл. 7). Это на два порядка меньше, чем в случае A . В согласии с экспериментом трехмерные расчеты показывают [135], что случаи B и C (но не A) больше соответствуют цинковой обманке. С точки зрения плотности состояний (правая часть фиг. 30) поверхностные уровни обычно сливаются с зоной проводимости, однако этот эффект может быть кажущимся, так как дискретные уровни никогда не пересекаются с краем зоны в k -пространстве.

ПОЛУКЛАССИЧЕСКИЕ МЕТОДЫ

Как было уже указано, проблема поверхностных состояний крайне трудна при решении ее методами квантовой механики в трехмерном случае. Желательно поэтому было бы ввести простые полуклассические аналоги задачи. Несмотря на их приближенный характер по сравнению с соответствующими квантовомеханическими моделями, преимущество и польза от применения таких моделей состоит в их простоте, а также тесной связи с известными представлениями структурной химии и классической электростатики. Кроме того, некоторые интегралы, необходимые при таком квантовомеханическом подходе, могут быть оценены, как будет показано ниже, с помощью этих полуклассических моделей.

Полуклассическим аналогом ковалентного кристалла служат тетраэдрические связи. В этом случае поверхность естественно представить с помощью оборванной или ненасыщенной связи. С другой стороны, полуклассической моделью ионного кристалла является цепочка ионов, причем поверхностным ионам отвечает уменьшенное координационное число, что обуславливает уменьшение поверхностного потенциала Маделунга. Очень кратко эти две, в некотором смысле альтернативные модели будут обсуждены позже. Конечно, для соединений типа $A^{III}B^{IV}$ и $A^{II}B^{VI}$ ни одна из моделей не является удовлетворительной, поэтому необходима определенная «интерполяция» между двумя крайними (ионным и ковалентным) случаями.

§ 9. Приближение ненасыщенных связей

Удобной моделью ковалентного кристалла является модель «шаров на стержнях». Свободная поверхность в такой модели представляется как плоскость ненасыщенных связей (их также называют нарушенными связями). Атомы в объеме имеют по четыре тетраэдрические

связи, занятые (ввиду наличия спина у электрона) двумя электронами, которые образуют заполненную оболочку из восьми электронов. Поверхностные атомы могут иметь одну или две оборванные связи (в зависимости от типа кристаллической плоскости). Три оборванные связи отвечают, вероятно, нестабильной конфигурации.

Обратимся теперь к рассмотрению некоторых геометрических соотношений в применении к структуре алмаза. Нас интересуют следующие свойства: тип кристаллической поверхности, число оборванных связей на атом, плотность упаковки атомов и плотность разорванных связей. Эти данные в случае Ge сведены в табл. 2.

Таблица 2

Идеализированные свойства неперестроенных поверхностей Ge *

Поверхность	Число разорванных связей на атом	Плотность поверхностных атомов	Число разорванных связей на 1 см ²	Относительное число разорванных связей на 1 см ²
(111)	1	$4/\sqrt{3} S^2 = 7,3 \cdot 10^{14} \text{ см}^{-2}$	$4/\sqrt{3} S^2$	1,0
(110)	1	$2\sqrt{2}/S^2 = 8,92 \cdot 10^{14} \text{ см}^{-2}$	$2\sqrt{2}/S^2$	1,23
(100)	2	$2/S^2 = 6,3 \cdot 10^{14} \text{ см}^{-2}$	$4/S^2$	1,73

* В действительности поверхности Ge и Si испытывают перестройку.

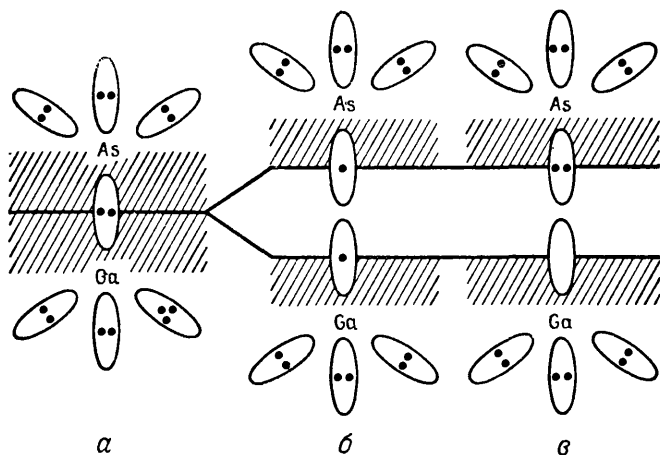
Пусть S — сторона элементарной ячейки гранецентрированного кубического кристалла ($S = 5,65 \text{ \AA}$ для Ge и $5,42 \text{ \AA}$ — для Si). Из табл. 2 видно, что наименьшей поверхностной плотностью оборванных связей обладает поверхность (111), несколько большей — (110) и самой большой — (100). Хандлер [136] использовал эти значения для интерпретации данных по поверхностным состояниям в Ge; Хафф и др. [29] применили их при описании поверхностных состояний в InSb; Аллен и Фаулер [137] учитывали их, сравнивая данные по величине работы выхода для различных поверхностей Ge. Основным недостатком приведенных в этой таблице сведений является то, что, согласно результатам ДЭНЭ, поверхности

испытывают перестройку, при этом ненасыщенные связи перераспределяются и частично насыщаются. Этот эффект будет рассмотрен в гл. 7, и в данной таблице он не учтен.

Концепция ненасыщенных связей была первым полуклассическим объяснением существования поверхностных состояний (первое квантовомеханическое объяснение было дано Таммом [14]). Как предполагалось, каждая оборванная связь отвечает одному или может быть нескольким электронным поверхностным состояниям. Хандлер [138] и Хафф с соавторами [29] предположили, что поверхностные состояния возмущают объемные состояния, формирующие валентную зону и зону проводимости. Однако попытки рассчитать энергетические уровни поверхностных состояний на основе представления о ненасыщенных связях сделано не было.

Ненасыщенные связи часто используются для объяснения *структурных* и *химических* свойств Si и Ge (см. § 17), а также свойств полярных поверхностей частично ионных полупроводников со структурой цинковой обманки и вюрцита. Рассмотрим, например, соединение GaAs, имеющее структуру цинковой обманки. В полярном направлении {111} кристалл состоит из чередующихся слоев атомов А и В. Идеальная плоскость А-типа (111) у GaAs образуется, если кристалл оканчивается плоскостью атомов галлия; идеальной плоскости В-типа (111) отвечает последняя плоскость атомов мышьяка. Экспериментально обнаружено, что плоскость А имеет иные по сравнению с плоскостью В свойства травления и окисления. Гато и Лавин [28, 139, 140] объяснили это следующим образом. Атомы в объеме GaAs, как показано на фиг. 31, а, имеют восемь электронов связи. После образования поверхности (111) путем разрушения одной связи (фиг. 31, б) может оказаться, что каждый из поверхностных атомов имеет по семь электронов. Однако, исходя из химических соображений, Гато и Лавин показали, что более вероятной является другая конфигурация, показанная на фиг. 31, в. В этом случае галлий (валентность 3) отдает свой валентный электрон атому мышьяка (валентность 5), принадлежащему другой свободной поверхности кристалла, дополняя его

оболочку до восьми электронов. Поверхностный атом Ga теперь имеет только шесть электронов связи. При этом будет проявляться тенденция к поверхностной sp^2 -гибридизации, и Ga станет химически относительно инертным. Напротив, поверхностные атомы As получают восемь электронов связи и поэтому будут стремиться сохранить sp^3 тетраэдрические связи. Ненасыщенная пара электронов у атома As делает его химически активным до-



Фиг. 31. Ненасыщенные связи на поверхности (111) GaAs.

нором, причем наиболее активным в окружении атомов окисла. Подобным же образом были объяснены [141] многие из наблюдаемых различий в химическом поведении А- и В-поверхностей. Аналогичные результаты были получены для полярных поверхностей А- и В-типов вюрцита. В рамках модели разорванных связей были интерпретированы [142] также результаты ранних работ по спонтанному изгибанию тонких пластинок (111) кристаллов соединений типа $A^{III}B^V$, однако впоследствии эти результаты были подвергнуты сомнению [143] ввиду того, что повреждения при полировке поверхности не были полностью удалены травлением и за изгиб были, вероятно, ответственны в основном остаточные напряже-

ния и (или) поверхностные окислы. Представление о ненасыщенных связях часто используют также в теориях зарождения и роста кристаллов [144].

Результаты, полученные в рамках метода ненасыщенных связей, оказались отчасти сомнительными после того, как эксперименты по ДЭНЭ показали, что предположительно идеальные полярные поверхности (111) цинковой обманки на самом деле перестраиваются [или, в некоторых случаях, образуют микрофасеточную структуру из неперестроенных плоскостей типа (110)]. Хейнеман [145, 146] предложил модель, в которой учитывалась информация, полученная из экспериментов с ДЭНЭ, допустив, что $1/4$ часть поверхностных атомов типа А несколько приподнята, а остальные $3/4$ опущены по отношению к А-плоскости (111) цинковой обманки. При этом предполагалось, что приподнятые атомы участвуют в основном в *s*-гибридизации, а опущенные — в *p*-гибридизации. Хейнеман [145, 146] и Хольт [147] считали заряд атомов А- и В-поверхностей равным $6 + 3/4$ и $6 + 5/4$ соответственно. Макрэй [148, 149] также использовал представление о ненасыщенных связях для объяснения ДЭНЭ на поверхности (110), а также на полярных поверхностях цинковой обманки.

В заключение можно, следовательно, отметить, что идея ненасыщенных связей оказалась полезной при изучении структуры и химических свойств чистых поверхностей, однако она совершенно неприменима для расчета энергетических соотношений и распределения собственных поверхностных состояний по энергиям, если не считать грубого предположения о том, что каждый ненасыщенной связи отвечает примерно одно поверхностное состояние.

§ 10. Метод потенциала Маделунга

Моделью ионного кристалла может служить упорядоченное расположение точечных зарядов. Хорошо известно, что на основе этой модели можно с успехом объяснить большинство структур ионных кристаллов [150] и энергий связи [151], причем это относится даже к тем системам, где некоторый вклад дает ковалентная связь.

Электронные уровни ионных и частично ионных кристаллов (NaCl , ZnO , ZnS) были рассчитаны полуклассическим методом Зейтцем [151], а также Моттом и Герни [152]. Они пользовались (вместо квантовой механики или метода ненасыщенных связей) обычной электростатикой и суммами Маделунга. Хотя вследствие электрон-электронного экранирования метод сумм Маделунга является идеализацией, он все же весьма полезен в применении к некоторым простым случаям [153].

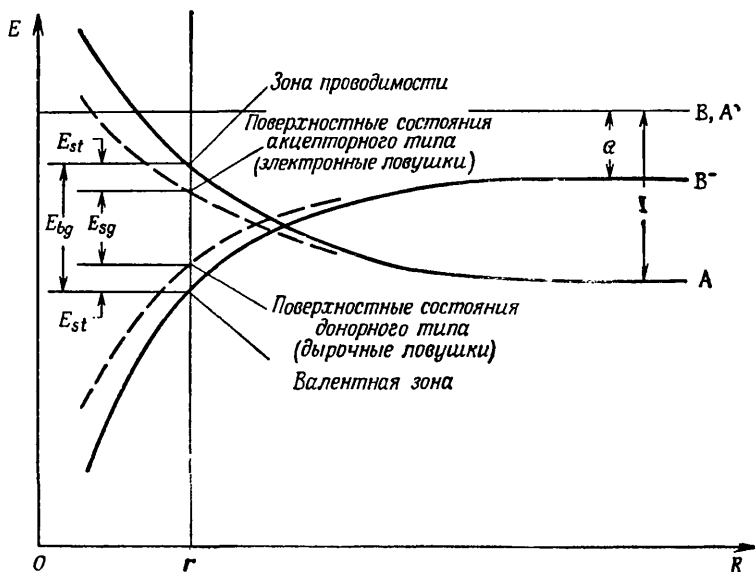
Левин и Марк [30] обобщили этот полуклассический метод на случай ионных поверхностных состояний. Хотя их подход является приближенным, он позволяет естественным образом классифицировать объемные и поверхностные состояния для всех ионных кристаллов и для всех типов ориентировки поверхностей. Для ковалентных и металлических кристаллов возможность подобной простой классификации до сих пор еще не была доказана. Метод потенциала Маделунга был впервые применен к объемным состояниям, а затем обобщен на поверхностные состояния. Основные численные результаты были получены в виде отношений параметров, характеризующих поверхностные и объемные свойства; предполагается, что отношения являются в данном случае более точной характеристикой, чем разности абсолютных величин.

Рассмотрим ионный кристалл типа АВ, где А обозначает катион, а В — анион. В объеме кристалла взаимодействие анионов приводит к образованию валентной зоны, заполненной при температуре 0 К. Взаимодействующие катионы формируют зону проводимости, которая при температуре 0 К свободна. В дальнейшем валентная зона и зона проводимости будут аппроксимированы дискретными уровнями и, таким образом, будет предполагаться, что связь имеет полностью ионный характер. Уточнения этой идеализированной картины будут рассмотрены позже.

Экспериментальное доказательство того, что в кристаллах типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$, $\text{A}^{\text{IV}}\text{B}^{\text{VI}}$ и $\text{A}^{\text{IV}}\text{B}^{\text{VII}}$ валентная зона имеет анионный характер, было дано Сванком [154] и будет нами рассмотрено в гл. 7. Построенный им для восемнадцати веществ график зависимости порога фотопроводимости от степени электроотрицательности

аниона дал четкую прямую линию. Эксперименты по ДЭНЭ также показывают, что в этих кристаллах ионная связь является доминирующей: электрически нейтральные (напоминающие шахматную доску) поверхности не испытывают перестройки (см. § 15).

Согласно Зейтцу [151], энергетические уровни сме-



Фиг. 32. Уровни энергии как функции межатомного расстояния R в ионных кристаллах типа АВ.

щаются с изменением межатомного расстояния R (параметр решетки) примерно так, как показано на фиг. 32. Когда $R \rightarrow \infty$, уровни изолированных атомов будут A и B^- . На фиг. 32 показана их связь с потенциалом ионизации I и сродством к электрону a . Так как низшим является уровень A , то стабильным состоянием разделенных атомов является *нейтральное* состояние. Если представить себе, что межатомное расстояние R монотонно убывает, то ввиду наличия потенциала Маделунга, симметричного относительно центра запрещенной зоны

$(I + a)/2$, уровни пересекаются. Для равновесного расстояния r самым низким оказывается уровень B^- , так что стабильное состояние ионного кристалла образуется из ионов A^+ и B^- . При этом ширина объемной запрещенной зоны (bulk bandgap) E_{bg} между валентной зоной и зоной проводимости равна

$$E_{bg} = 2V_b - I + a, \quad (10.1)$$

где V_b — потенциал Маделунга в объеме. Важно причем, что чем больше потенциал Маделунга, тем больше запрещенная зона. Подобным же образом [30] щель между зонами *поверхностных состояний* (surface state band gap) E_{sg} , показанная на фиг. 32, будет определяться соотношением

$$E_{sg} = 2V_s - I + a. \quad (10.2)$$

Здесь принимается, что параметры I и a имеют те же значения, что и в объеме, а поверхностный потенциал Маделунга равен V_s ($V_s \neq V_b$). Согласно фиг. 32, глубина ловушки, образованной поверхностным уровнем (surface state trap depth), E_{st} определяется выражением

$$E_{st} = \frac{E_{bg} - E_{sg}}{2}. \quad (10.3)$$

В этой простой электростатической модели *поверхностные состояния образуют симметричную пару состояний* A^- и B^+ -типа с одинаковым значением E_{st} . Формально величина E_{st} связана с разностью кулоновских интегралов $\alpha_A - \alpha'_A$ и $\alpha'_B - \alpha_B$, введенных в гл. 3. Так как отношения являются более точной характеристикой, чем абсолютные величины, можно переписать соотношения (10.1) — (10.3) в виде

$$\frac{E_{st}}{E_{bg}} = \frac{1 - \gamma}{2(1 - \mu)}. \quad (10.4)$$

Здесь параметр γ , описывающий геометрию поверхности, и параметр μ , характеризующий свойства ионов, определены следующим образом:

$$\gamma = \frac{V_s}{V_b}, \quad (10.5)$$

$$\mu = \frac{I - a}{2V_b}. \quad (10.6)$$

Основной общий вывод, который следует из соотношения (10.4), состоит в том, что поверхностные ловушки будут относительно глубже в кристаллах с *меньшим* γ и *большим* μ . Теперь займемся более подробным вычислением этих величин. При этом нормальными к поверхности смещениями поверхностных атомов при таком упрощенном рассмотрении можно пренебречь. Эти смещения составляют всего несколько процентов.

Потенциалы Маделунга, необходимые для вычисления параметра γ , определим следующим образом:

$$V_l = \frac{c_l e^2}{r}, \quad l = b \text{ или } s, \quad (10.7)$$

где $c_{b(s)}$ — объемная (поверхностная) константа Маделунга, e — единичный электрический заряд, z — валентность (возможно, дробная) и r — расстояние между анионом и катионом. Ввиду отсутствия достаточной информации величины r и z принимаются для объемных и поверхностных ионов одинаковыми. Значения коэффициента c_b для каменной соли, CsCl, цинковой обманки и вюрцита известны [155]: 1,748; 1,763; 1,638 и 1,641 соответственно. Эти величины получаются с помощью обычных кулоновских сумм

$$c_b = \sum_{i, j, k} \frac{q_{ijk}}{R_{ijk}} = \sum_{i, j, k} Q, \quad (10.8)$$

где $q_{ijk} = \pm 1$ в зависимости от знака заряда иона, помещенного в точке с индексами (i, j, k) , а R_{ijk} — расстояние узла (i, j, k) от начала координат $(0, 0, 0)$, отнесенное к расстоянию между ближайшими анионом и катионом [155]. Узел $(0, 0, 0)$ из суммы исключен. Для NaCl, например,

$$c_b = \left(\frac{6}{1} - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \frac{6}{\sqrt{4}} + \dots \right).$$

Аналогично *поверхностная* константа Маделунга может быть представлена [30] в виде

$$c_s = \sum_{i \geq 0, j, k} Q. \quad (10.9)$$

Здесь индекс i служит для нумерации узлов по нормали к поверхностной плоскости $i = 0$, а индексы j и k — для нумерации проекции узлов на эту плоскость. Например, для плоскости (100) NaCl в идеальном случае имеем

$$c_s = \left(\frac{5}{1} - \frac{8}{\sqrt{2}} + \frac{4}{\sqrt{3}} - \frac{5}{\sqrt{4}} + \dots \right).$$

Процедура суммирования этих плохо сходящихся рядов может быть заметно упрощена, если разложить величины c_b и c_s следующим образом:

$$c_b = \sum_{i=0} Q + \sum_{i>0} Q + \sum_{i<0} Q = \sum_{i=0} Q + 2 \sum_{i>0} Q, \quad (10.10)$$

$$c_s = \sum_{i=0} Q + \sum_{i>0} Q = \frac{1}{2} c_b + \frac{1}{2} \sum_{i=0} Q, \quad (10.11)$$

где суммирование по j и k производится точно так же, как и в формулах (10.8) и (10.9). Подобное разделение возможно лишь для некоторых ориентировок плоскостей поверхности, которые обладают соответствующей симметрией относительно плоскости $i \neq 0$. Например, плоскость (11 $\bar{2}$ 0) вюрцита отвечает этим требованиям, чего нельзя сказать о плоскости (10 $\bar{1}$ 0). Очевидно, сумма $\sum_{i=0} Q$ должна сходиться. Указанные условия не выполняются для всех слоистых поверхностей, таких, как (111) цинковой обманки, (0001) вюрцита и (111) каменной соли. Согласно экспериментальным данным по ДЭНЭ, эти слоистые поверхности часто оказываются так или иначе перестроенными; подробнее это будет рассмотрено в гл. 7. Таким образом, если разложения (10.10) и (10.11) возможны, то отношение $\gamma = c_s/c_b$ имеет вид

$$\gamma = \frac{1}{2} + \frac{1}{2} c_b^{-1} \sum_{i=0} Q, \quad (10.12)$$

т. е. параметр γ характеризует чисто геометрическое свойство кристаллической поверхности, так как он не зависит от природы составляющих кристалл анионов и катионов. Задача вычисления γ сводится теперь к вычислению константы Маделунга некоторой плоскости

(суммирование по j и k при $i = 0$), поскольку величина γ известна. Значения γ лежат в интервале $1/2 < \gamma < 1$. Они были рассчитаны [30] для разных электрически нейтральных кристаллических плоскостей. Например, для плоскости (100) каменной соли $\gamma = 0,96$. Большое значение γ показывает, что поверхностные состояния вблизи плоскости (100) каменной соли весьма похожи на ее объемные состояния. Два других примера: $\gamma = 0,92$ для плоскости (11 $\bar{2}$ 0) вюрцита и $\gamma = 0,91$ для плоскости (110) цинковой обманки. Таким образом, поверхностные состояния вюрцита и цинковой обманки должны быть расположены ближе к центру запрещенной зоны, чем поверхностные состояния каменной соли. Отношение ρ координационных чисел атомов у поверхности и в объеме также должно значительно отличаться в двух рассмотренных выше случаях. Действительно, в случае плоскости (100) каменной соли $\rho = 5/6$, а для (11 $\bar{2}$ 0) вюрцита и (110) цинковой обманки $\rho = 3/4$. Таким образом, должно наблюдаться большое различие между ионами на поверхности и в объеме для вюрцита и цинковой обманки, чего нельзя ожидать в случае каменной соли. Использование счетных машин позволяет, вообще говоря, выполнить суммирование в выражении (10.9) даже тогда, когда разложения (10.10) и (10.11) несправедливы; это было сделано Носкером в случае поверхности (10 $\bar{1}$ 0) вюрцита. В результате было получено значение $\gamma = 0,88^1$).

Вычисление параметра μ в (10.6) требует знания величин I , a , z , γ и r . Поскольку энергии I и a сами зависят от дробной валентности z , параметр μ можно представить в виде $\mu(z)$, вычислить для целых значений, пользуясь первым и вторым ионизационными потенциалами и сродством к электрону, а затем интерполировать между этими значениями. Рассмотрим, например, кристалл CdS. Взяв $z = 2$ и учитывая значения $I(2)$ и $a(2)$, получим $\mu(2) = 0,58$. Когда $z = 1$, $I(1)$ и $a(1)$, имеем $\mu(1) = 0,41$. С помощью линейной

¹) Рассчитанные Носкером с помощью вычислительной машины значения γ [опубликовано в Surface Science, 20, 421 (1970)], вероятно, более точны, чем полученные Левином и Марком [30].

экстраполяции к предложенному Бирманом значению дробной валентности [156] $z \sim 0,5$, получим $\mu(0,5) = 0,33$. Левин и Марк [30] вычислили значения μ для 46 галоидов, окислов и сульфидов. В целом с увеличением степени ионности соединения величина μ уменьшается. Например, для NaCl ($z = 1$) имеем $\mu(1) = 0,06$. В принципе ничто не препятствует тому, чтобы отношение E_{st}/E_{bg} в (10.4) стало больше единицы, особенно в случае больших значений μ . Это может иметь место, когда валентность очень мала ($z \sim 0,1$), например у некоторых соединений типа $A^{III}B^V$ с узкой запрещенной зоной ($E_{sg} < 1$ эВ).

Мы здесь приближенно приняли, что валентная зона и зона проводимости являются дискретными уровнями. На самом деле и объемные, и поверхностные состояния должны расщепиться в зоны. Часть зоны поверхностных состояний, вероятно, перекрывается с объемными зонами и окажется за пределами запрещенной зоны; эту часть распределения поверхностных состояний наблюдать будет довольно трудно. (См. гл. 7, в особенности фиг. 34, д.)

Исторически полуклассический метод потенциала Маделунга был первым методом, выявившим определенную симметрию между поверхностными состояниями А- и В-типа. Очевидным недостатком этого метода было то, что он не основывался на решении уравнения Шредингера. По этой причине затем были предприняты и квантовомеханические расчеты. Прямой метод кристаллических орбиталей или метод края зоны ЛКАО (§ 1) и метод Матье показали, что ионные поверхностные состояния имеют довольно общую природу и возникают независимо от используемого метода расчета.

Экспериментальные результаты, относящиеся к ионным поверхностным состояниям в соединениях элементов II—VI и III—V групп, рассмотрены в § 17. Там показано, что данные эксперимента весьма удовлетворительно согласуются с теорией ионных поверхностных состояний, которая была здесь развита. Марк обобщил рассмотренный формализм на случай изучения строения кристалла [122] и хемосорбционных состояний [158].

СРАВНЕНИЕ ТЕОРИИ И ЭКСПЕРИМЕНТА

Именно за последние десять лет было собрано огромное количество экспериментальных данных по поверхностным состояниям у атомно чистых поверхностей. Основной причиной такого прогресса явилась возможность получения на установках условий сверхвысокого вакуума (10^{-11} — 10^{-10} мм рт. ст.). При давлении 10^{-6} мм рт. ст. на поверхности примерно за 1 с образуется одноатомный слой остаточных газов, однако под давлением 10^{-10} мм рт. ст. этот промежуток времени составляет уже около 10^4 с или более одного часа. Таким образом, эксперименты в условиях сверхвысокого вакуума могут выполняться с минимальной опасностью загрязнения. Для некоторых сочетаний адсорбента и адсорбата коэффициент прилипания имеет низкие значения, что помогает сохранить поверхность чистой в течение длительного времени.

В условиях сверхвысокого вакуума удастся наблюдать *собственные состояния* на свободной поверхности в противоположность *несобственным состояниям*, обусловленным присутствием посторонних атомов. Источниками несобственных состояний могут служить химически и физически адсорбированные атомы, поверхностные окислы, остатки электролитов при травлении, осажденные металлы. В дальнейшем мы подробно рассмотрим важнейшие эксперименты и полученные результаты, касающиеся собственных поверхностных состояний, которые являются главным предметом нашего обзора. Несобственных состояний мы коснемся очень кратко в п. 4 § 17. Некоторые осложнения, обусловленные иногда процессом приготовления поверхности, такие, как *p*-инверсия в результате дегазации стенок, будут рассмотрены особо в § 17. Другие факторы, иногда усложняющие картину, — это *краевые состояния*, вызванные поверхностными дислокациями или ступеньками скола, а также

атомная перестройка. Они будут также обсуждены отдельно в § 15 и 16.

Металлы, изоляторы и полупроводники обладают свободными поверхностями и, следовательно, могут в принципе иметь поверхностные состояния¹⁾. Можно ли их каким-либо образом наблюдать экспериментально? В металлах энергетическая щель вблизи уровня Ферми отсутствует, и плотность электронов в объемных состояниях в этой области высока. Поэтому металлические поверхностные состояния трудно экспериментально отделить от объемных. Однако металлические поверхности обладают другими характерными свойствами, которые могут быть обнаружены. Сюда относятся, в частности: 1) поверхностные плазменные потери [159, 160]; 2) поверхностные парамагнитные свойства [161]; 3) характеристики рассеяния на поверхности [162, 163]; 4) поведение волновой функции, определяемое методами Оже [164]; 5) ненасыщенные связи, которые проявляются в яркости фасеточных граней при эмиссии из металлов во внешнем поле [165, 166], и 6) участие в хемосорбции [167, 168]. Однако ни одно из перечисленных свойств, по-видимому, не может быть непосредственно связано с собственными поверхностными состояниями, которые теоретически рассматривались в предыдущих параграфах. Поэтому в дальнейшем мы не будем касаться этих свойств.

Изоляторы, такие, как щелочно-галогидные соединения, галогиды серебра, окись алюминия и стекла, часто насыщены в объеме электронными ловушками (как для электронов, так и для дырок). В щелочно-галогидных соединениях эти ловушки имеют вполне определенные свойства и называются *F*-центрами, *U*-центрами, *V*-центрами и т. д. При этом каждое название соответствует определенному типу точечных дефектов или дефектов, связанных с атомными скоплениями. Таким образом, ловушки в объеме и поверхностные состояния, если они имеются, далеко не так просто различить. Кроме того, в таких материалах ионная проводимость часто

¹⁾ См. примечание 2 на стр. 114. — *Прим. перев.*

превалирует над электронной. Эти трудности не позволяют до сих пор получать данные по собственным поверхностным состояниям в изоляторах с большой запрещенной зоной. Для AgCl были получены некоторые результаты [169], однако не в условиях сверхвысокого вакуума. Контактные состояния типа центров окраски в галоидах серебра интенсивнее всего изучались при исследовании фотографических процессов [170].

Полупроводники, по-видимому, обладают самыми подходящими свойствами для наблюдения поверхностных (и контактных) состояний. Поверхностное состояние (если оно имеется), находящееся в запрещенной зоне, можно легко выявить. В отдельных случаях электронные измерения позволяют получить разрешения до 10^{10} состояний на 1 см^2 (п. 4 § 17), что соответствует плотности 10^{-5} монослоя. Полупроводниковая технология позволяет выращивать кристаллы с относительно малым числом объемных ловушек. Поэтому реакция поверхностных состояний на внешние поля, облучение, окружающую среду и другие воздействия может быть определена относительно легко. Ясно поэтому, что теория поверхностных состояний и эксперимент лучше всего могут быть сопоставлены именно для полупроводников. Для Si , Ge и соединений элементов III—V и II—VI групп это сопоставление проводится в п. 4 § 17. Теория поверхностных состояний и эксперимент, по-видимому, находятся в лучшем согласии для неперестроенных ионных поверхностей полупроводников типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$ и $\text{A}^{\text{II}}\text{B}^{\text{VI}}$, чем для поверхностей Si и Ge , испытывающих перестройку.

§ 11. Основные явления

Чтобы понять строение полупроводниковых поверхностей, необходимо с физической точки зрения рассмотреть явления в слое объемного заряда. Ввиду недостатка свободных электронов и дырок в объеме полупроводника присутствие небольшого числа поверхностных состояний может определить поведение этого слоя. Подобный «эффект усиления» делает возможным простое детектирование поверхностных состояний в полупроводниках.

Необходимость появления слоя объемного заряда следует из решения трехмерного уравнения Пуассона

$$\nabla^2 V(x, y, z) + \frac{\rho(x, y, z)}{K\epsilon_0} = 0, \quad (11.1)$$

где ∇ — локальный потенциал (в вольтах), K — локальная диэлектрическая постоянная, ϵ_0 — диэлектрическая постоянная вакуума ($8,85 \cdot 10^{-12}$ Ф/м)¹⁾ и ρ — плотность свободных и захваченных в ловушки зарядов (Кл/м³). Поверхностные состояния можно учесть в выражении для ρ , вводя соответствующие δ -функции. Конечно, если $\rho = 0$, неоднородности заряда отсутствуют, а в результате отсутствуют и поверхностные состояния. Хотя определенный прогресс был достигнут в рамках трехмерного варианта уравнения (11.1), однако для изучения объемных [171] и поверхностных состояний [172] часто используют вместо (11.1) идеализированную *одномерную модель*. При этом роль направлений, параллельных свободной поверхности, считается несущественной. Такая идеализация прежде всего исключает все многочастичные эффекты взаимодействия [173]²⁾ между электронами в поверхностных состояниях. Кроме того, принимается, что в объеме кристалл является однородным, даже несмотря на то, что объемные доноры и акцепторы являются точечными центрами. Далее, эта модель постулирует, что положения уровней относительно краев зоны не зависят от V или от заполнения поверхностного состояния. Другими словами, уровни смещаются вместе с зонами. Для объемных состояний такое приближение было ранее обосновано, однако оно может оказаться несправедливым для поверхностных состояний. Дело в том, что нескомпенсированный заряд поверхностных состояний может вызвать заметное макроскопическое поле в объеме. Поэтому поверхностные состояния могут в этом случае играть более существенную роль, чем такое же число хаотично распределенных объемных состояний.

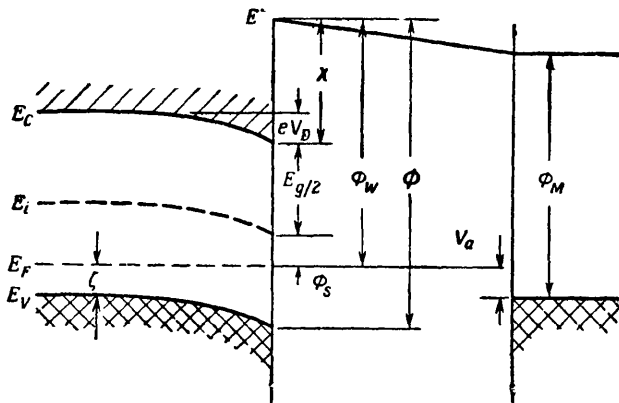
¹⁾ Авторы пользуются системой единиц МКСМ. — *Прим. перев.*

²⁾ Обсуждению роли многочастичных, корреляционных эффектов при анализе данных по электронному парамагнитному резонансу в полупроводниках посвящена работа [332]. В этой работе температурная зависимость сигнала ЭПР объяснена на основе предположения о существовании поверхностных состояний. — *Прим. перев.*

Статистическое заполнение поверхностных состояний принимается в настоящее время таким же, как в ферми-статистике объемных состояний. Однако, согласно Хейнеману [174], это предположение, возможно, будет изменено в будущем из-за кулоновского взаимодействия электронов в поверхностных состояниях. Наконец, потенциал изображения у свободной поверхности полупроводника никогда точно не рассматривался в рамках традиционного решения уравнения (11.1). Обычно подразумевается, что этот потенциал учтен (весьма грубо) в величине сродства к электрону всего кристалла χ как граничное условие. Это упущение не очень существенно для объемных состояний где-то вблизи поверхности, однако оно приобретает серьезное значение, когда речь идет о поверхностных состояниях.

Решения уравнения (11.1), получение которых даже при описанных выше упрощениях сопряжено подчас с весьма трудными выкладками, в настоящее время хорошо известны. За математическими деталями читатель отсылается к работам Мейни и др. [7] и Франкля [8]. В этих работах имеются соотношения и графики, связывающие все наиболее важные переменные, и в частности потенциалы и плотности заряда. В этом обзоре мы не будем выводить или воспроизводить эти уравнения и зависимости. Будет введена только терминология и отмечены наиболее показательные качественные результаты. Они помогут читателю понять, *что* именно измеряется и *как* полученные данные интерпретируются экспериментаторами. Комментарии к некоторым экспериментальным работам (до 1966 г.) даны вышеупомянутыми авторами [7, 8]. Эти комментарии полезны, но часто загромождены огромным количеством деталей, которые могут быть несущественными для теоретика или экспериментатора, интересующегося распределением собственных поверхностных состояний. Цель данного обзора, — опуская большинство деталей, выделить существенные черты экспериментально наблюдаемого распределения поверхностных состояний и включить ссылки на последние работы (вплоть до середины 1969 г.). Подробности читатель сможет найти в оригинальных статьях.

Основные энергетические величины, используемые экспериментаторами для описания свойств поверхностных состояний, схематически показаны на фиг. 33. Эта схема отвечает кристаллу p -типа ($p_b > n_b$) с истощенным слоем ($p_b > p_s$ и $p_s > n_s$). Здесь p_b , n_b , p_s и n_s обозначают соответственно объемную плотность дырок, объемную плотность электронов, поверхностную плотность дырок и поверхностную плотность электронов.



Фиг. 33. Параметры, описывающие слой пространственного заряда в полупроводнике p -типа.

При отсутствии приложенного внешнего поля и поверхностных состояний искривление зон eV_D пропадает.

Можно также рассмотреть сходные диаграммы для кристалла p -типа с обогащенным слоем ($p_b < p_s$ и $p_s > n_s$) или для кристалла p -типа с инверсным слоем ($p_b > p_s$, $p_s < n_s$). Также может быть рассмотрен аналог кристалла n -типа с истощенным, инверсным или обогащенным слоями. Определения энергий и соотношения, описанные ниже, следуют из фиг. 33. Они в незначительной степени зависят от выбора того или иного из перечисленных выше случаев. На диаграмме показан также плоский металлический электрод, находящийся справа от полупроводника. Если этот электрод действительно имеется, то с его помощью к поверхности полупроводника может быть приложено внешнее поле. Предполагает-

ся, что при изучении чистых поверхностей полупроводник и пластину (электрод) разделяет вакуумный промежуток (в случае контактных состояний это может быть слой окисла). Характерные энергии следующие: вакуумный уровень отсчета E' ; край зоны проводимости E_c ; край валентной зоны E_v ; собственный уровень $E_i \approx \frac{1}{2}(E_c + E_v)$; энергетическая щель $E_g = E_c - E_v$; энергия Ферми E_F ; работа выхода $\phi_w = E' - E_F$; фотоэмиссионный порог (экстраполированный кубической зависимостью) Φ , который, как это, в частности, изображено, зачастую соответствует краю валентной зоны у поверхности (точнее следовало бы назвать Φ кристаллическим потенциалом ионизации); сродство к электрону χ ; параметр легирования $\xi = E_F - E_v$; искривление зоны eV_D ; поверхностный потенциал ϕ_s ; работа выхода плоского металлического электрода ϕ_M и приложенное напряжение V_a . Параметр легирования для несобственного полупроводника p -типа приблизительно равен

$$\xi = kT \ln \frac{N_v}{N_A}, \quad (11.2)$$

где N_A — плотность акцепторов (см^{-3}), а N_v — эффективная плотность состояний валентной зоны (для состояний, лежащих в интервале $\sim kT$ вблизи края зоны). Как обычно, k — постоянная Больцмана и T — температура. Выпишем некоторые соотношения между указанными энергетическими параметрами для образца p -типа (считая, что Φ отвечает краю валентной зоны):

$$\Phi = \chi + E_g, \quad (11.3)$$

$$\phi_w = \chi - eV_D + E_g - \xi, \quad (11.4)$$

$$\phi_s = \frac{1}{2} E_g - eV_D - \xi. \quad (11.5)$$

В соответствии с фиг. 33 в равенствах (11.3) — (11.5) потенциал V_D считается положительной величиной. В отсутствие поверхностных состояний и без приложенного внешнего поля $V_D = 0$ и зоны остаются плоскими вплоть до самой поверхности. В случае измерения фотоэмиссии с поверхностных состояний Φ более не соответствует краю валентной зоны, и вместо (11.3) имеем $\Phi \leq \chi + E_g$.

Известно, что величина χ зависит от геометрии кристаллической поверхности [137]. Высказывались даже соображения, что для сильно вырожденного кремния p -типа χ зависит, кроме того, от типа и степени легирования [175—177]. Тем не менее соотношения (11.3) — (11.5) часто используют по крайней мере для предварительного рассмотрения поверхностных состояний. Исключая из (11.3) и (11.4) параметр χ , получаем

$$eV_D = \Phi - \phi_w - \xi. \quad (11.6)$$

Таким образом, измеряя порог фотоэмиссии, работу выхода и степень легирования, можно получить изгиб зоны в отсутствие внешних полей, т. е. величину, подчиненную указанным выше условиям. Разумеется, для образцов n -типа выписанные уравнения должны быть изменены в соответствии с тем, что ξ в образцах n -типа отсчитывается от края зоны проводимости. Замена $\xi \rightarrow (E_g - \xi)$ в (11.3) — (11.6) дает соответствующие соотношения в случае полупроводников n -типа. (При определении величины Φ по току фотоэмиссии как функции энергии фотонов имеются некоторые тонкости; их мы рассмотрим в п. 3 § 13.)

Если $V_D \neq 0$, это означает, что имеются поверхностные состояния. Различные способы вычисления V_D , кроме основанного на выражении (11.6), приведены ниже. Самым непосредственным методом определения V_D являются эксперименты с фото-э. д. с. При этом свет энергии $h\nu > E_g$ и интенсивности I вызывает образование высокой концентрации свободных электроннодырочных пар у поверхности. Это приводит к выравниванию формы зон. Обычно строят график ΔV как функции I (с помощью метода Кельвина или измерений емкости). При насыщении $V = V_D$. При этом необходимо вычесть поле Дембера, возникающее из-за того, что электроны и дырки, созданные светом у поверхности и обладающие различными подвижностями, диффундируют в объем.

Другой способ получения V_D состоит в измерении проводимости вдоль поверхности. Здесь главным является предположение о том, что заряд в поверхностных состояниях относительно малоподвижен и проводи-

мость в основном обеспечивается объемной областью. Это подтверждается другими экспериментами на Si и Ge [178, 179]. Кристалл может быть сколот в сверхвысоком вакууме. Если $\Delta\sigma$ (См/ед. площади) есть изменение поверхностной проводимости на единицу площади, произошедшее в результате скола, то изменение пространственного заряда у поверхности $\Delta p/e$ (см⁻²) можно оценить с помощью уравнения, которое в простейшем виде (проводимость p -типа) будет выглядеть следующим образом:

$$\Delta\sigma = e\Delta p\mu_{ps}, \quad (11.7)$$

где μ_{ps} [см²/ (В·с)] — подвижность дырок у поверхности. Вообще говоря, поверхностная подвижность меньше, чем объемная, так что при вычислении μ_{ps} необходимо оценить вклад соответствующего механизма рассеяния (зеркального, диффузного или смешанного). Наиболее широко применимое соотношение принадлежит Шрифферу [180]¹⁾. Поскольку Δp известно из (11.7), V_D находится с помощью стандартного соотношения между Δp и V_D . Дело обычно осложняется введением в (11.7) слагаемого $\Delta p\mu_{ps}$, связанного с изменением электронной проводимости у поверхности, и введением квантованных «легких дырок», если имеется сильно инверсный слой [181].

Если $\Delta\sigma$ не удастся получить в процессе скола, как описано выше, то V_D все же можно найти, но менее удовлетворительным и более косвенным путем. Предположим, что кристалл расколот, но его объемная проводимость не известна ввиду неопределенности состава или геометрии кристалла. Тогда можно сразу сказать, какая часть проводимости σ обусловлена поверхностным слоем ($\Delta\sigma$), а какая — объемом (σ_0), поскольку проводящие пути действуют параллельно. Объемный вклад может тогда быть исключен, если известна величина $\sigma_{\min}$ — минимальная проводимость, которая получается после раскочки зон приложенным внешним полем или после добавления адсорбата p - или n -типа. Величина σ — $\sigma_{\min}$ связывается затем известным образом с V_D [8].

¹⁾ См. также [8], стр. 110.

Выше мы остановились на важной роли изгиба зоны V_D и способах его измерения. Вообще говоря, если измерения на чистой поверхности дают изгиб зон в отсутствие адсорбции и внешних полей, то *должны* существовать поверхностные состояния. Чтобы определить, сколько имеется поверхностных состояний и как распределены их энергии, требуется более сложный анализ и экспериментальная техника. На этих вопросах мы теперь кратко и остановимся.

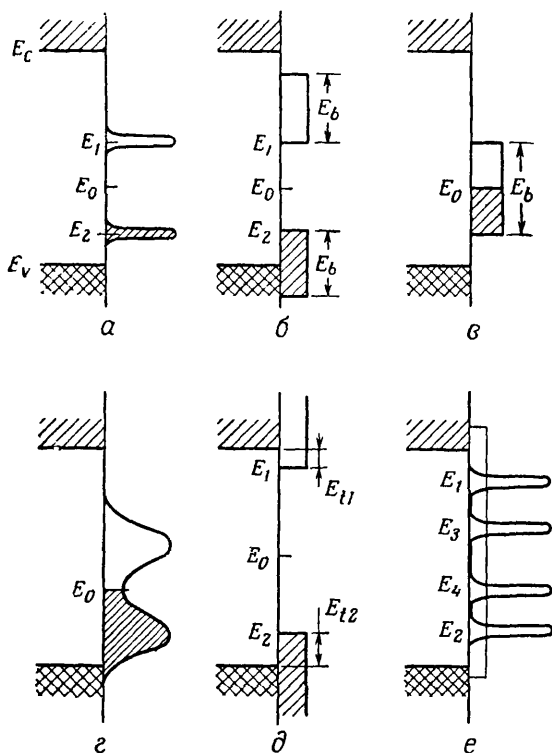
§ 12. Распределение поверхностных состояний

Если $V_D \neq 0$, то должны существовать поверхностные состояния. Как с помощью экспериментальных данных получить распределение поверхностных состояний? Из-за особенностей статистики Ферми окончательного ответа на этот вопрос не существует. Легко поддаются изучению только те состояния, которые находятся на самой поверхности Ферми или где-то вблизи нее. Например, электронные состояния, расположенные на несколько электрон-вольт выше зоны проводимости металла или полупроводника, могут быть исследованы только с применением экспериментов по поглощению света или по рассеянию горячих электронов. Таким образом, в целом вопрос остается открытым. Если бы удалось «пройти» уровнем Ферми по всей существенной области энергетической зонной структуры и если бы можно было поставить эксперименты в широком интервале температур, проблема была бы решена. Структуру поверхностных состояний можно было бы тогда определить однозначно. Именно потому, что на практике удается сместить уровень Ферми только на небольшую величину, а температурный интервал узок, полученные данные ограничены и полученная информация, касающаяся поверхностных состояний, неполна. Другим следствием малой ширины слоя, который может быть пройден уровнем Ферми, является то, что полное число (заряженных и нейтральных) поверхностных состояний часто приходится характеризовать его нижним пределом. Чтобы пояснить эти моменты, рассмотрим теперь некоторые полезные модели поверхностных состояний.

Аналогично объемным примесным состояниям поверхностные состояния можно рассматривать либо как «акцептороподобные», либо как «донороподобные». По условию, акцептороподобные состояния нейтральны, если они свободны, и отрицательно заряжены, если заполнены одним электроном. Донороподобные состояния заряжены положительно, когда пусты, и нейтральны, когда они заняты одним электроном. Для полупроводников, более близких к диэлектрику ($E_g > 2$ эВ), терминология несколько отличается [30]. Акцептороподобные состояния эквивалентны электронной ловушке (нейтральной в отсутствие электронов и заряженной отрицательно при наличии одного электрона). Донороподобные состояния эквивалентны дырочной ловушке (нейтральной, когда в ней нет дырки, и заряженной положительно, когда одна дырка захвачена). Термины «донор и акцептор» и «электронная и дырочная ловушки» будут использоваться в соответствующих ситуациях. Для поверхностных состояний, по-видимому, характерно такое положение, когда акцепторные уровни (электронные ловушки) оказываются расположенными *выше* донорных (дырочные ловушки). В этом отношении они напоминают поведение ловушек в кристаллах с лучшими изолирующими свойствами [182]. Тем, кто знаком только с мелкими примесными донорами и акцепторами в Si и Ge, такое поведение может показаться необычным. В самом деле, мелкие доноры лежат чуть ниже зоны проводимости, а мелкие акцепторы — чуть выше валентной зоны. С собственными поверхностными состояниями (и ловушками в объеме) дело обстоит как раз наоборот. Акцепторные (донорные) состояния обычно лежат ниже (выше) зоны проводимости (валентной зоны). Из одного лишь этого факта можно заключить, что поверхностные состояния *не* похожи на мелкие доноры и акцепторы. (Как указано выше, в этом параграфе рассматриваются только *собственные* поверхностные состояния. Высказанное утверждение, конечно, изменится, если присутствуют также *несобственные* поверхностные состояния.)

Может случиться, что некоторые из поверхностных состояний окажутся вне запрещенной зоны. Такой случай уже был рассмотрен в § 8, где было показано, что

он возможен даже тогда, когда дисперсионные кривые поверхностных и объемных состояний в k -пространстве не пересекаются.



Фиг. 34. Некоторые полезные модели собственных поверхностных состояний.

a — два дискретных уровня; b — две зоны; c — слившиеся зоны; $г$ — ch -подобные зоны; $д$ — зоны состояний ионного типа; e — возможное распределение несобственных состояний.

В качестве простого примера распределения поверхностных состояний рассмотрим показанную на фиг. 34, a модель *двух дискретных уровней*. Пусть $N_1(N_2)$ — число доступных поверхностных состояний акцепторного (донорного) типа, а $-Q_1(Q_2)$ — отрицательный (положительный) заряд, связанный с заполненными (пустыми)

акцепторными (донорными) состояниями. Соответствующие энергии — E_1 и E_2 . Полный заряд поверхностных состояний $Q_{ss} = Q_1 + Q_2$ с учетом ферми-статистики и единичного статистического веса будет равен

$$Q_{ss} = -eN_1 \left(1 + \exp \frac{E_1 - E_F}{kT}\right)^{-1} + eN_2 \left(1 + \exp \frac{E_F - E_2}{kT}\right)^{-1}. \quad (12.1)$$

Примем для простоты, что $N_1 = N_2 = N$, и пусть по определению

$$E_0 = \frac{1}{2} (E_1 + E_2), \quad \Delta E_F = E_F - E_0, \quad (12.2)$$

$$\frac{1}{2} E_{sg} = E_1 - E_0 = E_0 - E_2.$$

Новые величины E_0 и ΔE_F определяют соответственно энергию нейтрального состояния, когда $Q_{ss} = 0$, и положение энергии Ферми относительно этого уровня. Вместо (12.1) теперь получим

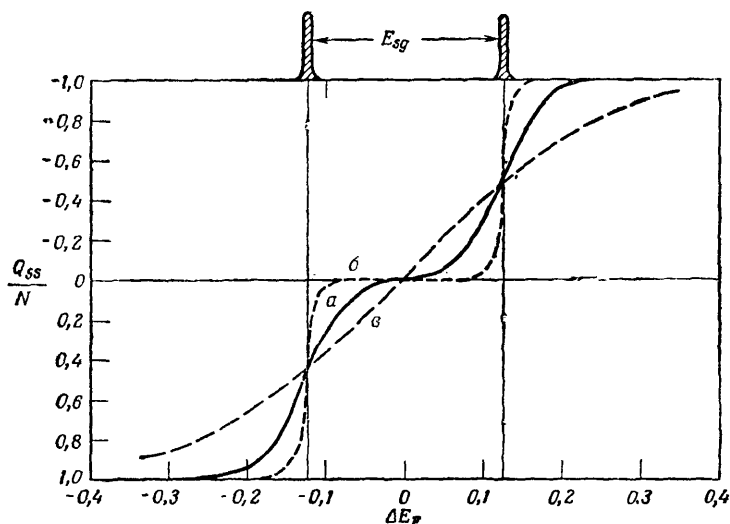
$$-\frac{Q_{ss}}{eN} = \left(1 + \exp \frac{1/2 E_{sg} - \Delta E_F}{kT}\right)^{-1} - \left(1 + \exp \frac{1/2 E_{sg} + \Delta E_F}{kT}\right)^{-1}, \quad (12.3)$$

что можно представить в виде

$$-\frac{Q_{ss}}{eN} = \frac{\text{sh}(\Delta E_F/kT)}{\text{ch}(E_{sg}/2kT) + \text{ch}(\Delta E_F/kT)}. \quad (12.4)$$

Функция $-Q_{ss}/eN$ описывает нормированный отрицательный заряд поверхностных состояний. Когда $\Delta E_F > < 1/2 E_{sg}$, эта функция напоминает гиперболический синус, а при $\Delta E_F > 1/2 E_{sg}$ стремится к насыщению. График зависимости Q_{ss}/N от ΔE_F показан на фиг. 35 для $T = 75, 300$ и 1200 К. Энергетическая щель между поверхностными состояниями E_{sg} считается (для простоты) не зависящей от температуры. На фигуре она выбрана равной $0,25$ эВ. Такое значение, по-видимому, характерно для Si. Замечая, что величины ΔE_F и $1/2 E_{sg}$ входят

в выражения вместе с температурой всегда в виде отношений $\Delta E_F/kT$ и $1/2 E_{sg}/kT$, можно унифицировать характерные температуры и энергии. Таким образом, энергетическую щель можно записать через температуру, т. е. величина E_{sg} равна $40 kT$ при 75 К , $10 kT$ — при комнатной температуре (300 К) и $2,5 kT$ — при 1200 К .



Фиг. 35. Модель двух дискретных уровней. Заряд поверхностных состояний как функция положения уровня Ферми и температуры. $a - T=300 \text{ К}$; $б - T=75 \text{ К}$; $в - T=1200 \text{ К}$.

Рассмотрение фиг. 35 показывает, что даже в случае простой модели распределения (пара дискретных уровней) имеется заметное изменение измеряемой величины заряда Q_{ss} как функции энергии и температуры. В конечном счете это обусловлено проникновением максвелловского экспоненциального «хвоста» за энергетический уровень поверхностного состояния. В пределе высоких температур Q_{ss} меняется линейно с ΔE_F и, по всей видимости, не зависит от дискретного характера поверхностных состояний, которые ответственны за воз-

никающий заряд. Наклон в точке E_0 растет с температурой; эту зависимость можно получить из (12.4):

$$-\left(\frac{\partial Q_{ss}}{\partial \Delta E_F}\right)_{E_0} = \frac{eN/kT}{1 + \text{ch}(E_{sg}/2kT)}. \quad (12.5)$$

Величина наклона тесно связана с «подвижностью в эффекте поля». Об этом речь будет идти ниже. Фиг. 35 отличается от фиг. 5.4 в работе Мейни с соавторами [7], где по оси абсцисс отложены *приведенные* энергии вида E/kT . В последнем случае трудно выделить температурную зависимость.

В качестве второго простого примера рассмотрим *двухзонную* модель поверхностных состояний, показанную на фиг. 34, б. Примем, что обеим зонам отвечает *не зависящая от энергии плотность* состояний $\bar{N}$ ($\text{см}^{-2} \cdot \text{эВ}^{-1}$). В таком случае имеем

$$Q_{ss} = -e\bar{N} \int_{E_1}^{E_1+E_b} \left(1 + \exp \frac{E-E_F}{kT}\right)^{-1} dE + e\bar{N} \times \\ \times \int_{E_2-E_b}^{E_2} \left(1 + \exp \frac{E_F-E}{kT}\right)^{-1} dE. \quad (12.6)$$

Здесь E_b — «ширина» акцепторной и донорной зон. (Для простоты принято, что $\bar{N}_1 = \bar{N}_2 = \bar{N}$ и $E_{b1} = E_{b2} = E_b$.) Допустим также для простоты, что величину $\bar{N}$ можно оценить с помощью модели двумерного газа свободных электронов. Электроны двукратно вырождены по спину и характеризуются некоторой подходящей изотропной эффективной массой m_s^* . В этом случае

$$\bar{N} = 4\pi m_s^* h^{-2} = 4,2 \cdot 10^{14} \frac{m_s^*}{m_0} \text{см}^{-2} \cdot \text{эВ}^{-1}. \quad (12.7)$$

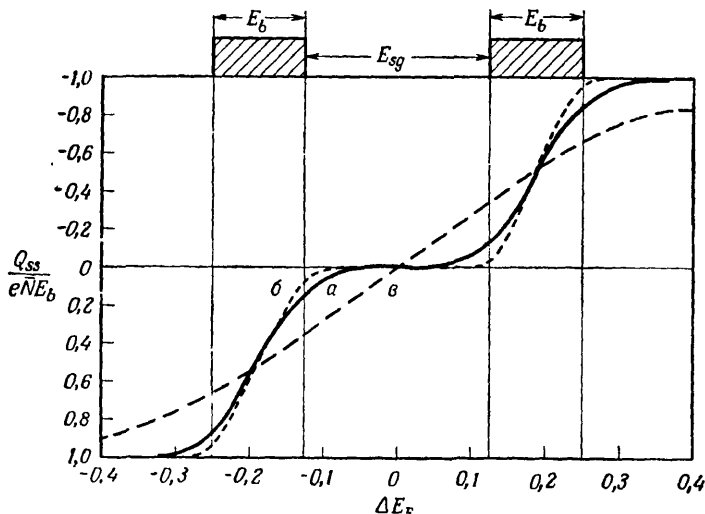
Приближение (12.7) должно, очевидно, нарушаться при отходе от края зоны поверхностных состояний. Однако принято считать, что оно применимо для большинства зон поверхностных состояний.

В противоположность трехмерной зонной задаче интеграл в формуле (12.6) удастся взять точно.

Получаем

$$\begin{aligned} \frac{Q_{ss}}{e\bar{N}E_b} = & \left(\frac{kT}{E_b}\right) \ln \left\{ \left[1 + \exp\left(\frac{1}{2}E_{sg} + E_b - \Delta E_F\right)/kT \right] \times \right. \\ & \times \left. \left[1 + \exp\left(\frac{1}{2}E_{sg} + \Delta E_F\right)/kT \right] \right\} - \\ & - \left(\frac{kT}{E_b}\right) \ln \left\{ \left[1 + \exp\left(\frac{1}{2}E_{sg} - \Delta E_F\right)/kT \right] \times \right. \\ & \times \left. \left[1 + \exp\left(\frac{1}{2}E_{sg} + E_b + \Delta E_F\right)/kT \right] \right\}. \quad (12.8) \end{aligned}$$

Эта зависимость изображена на фиг. 36 для значений $E_b = E_{sg} = 0,25$ эВ для трех температур: 75, 300 и 1200 К. Как можно видеть, эта функция обладает следующими свойствами. Если уровень Ферми расположен



Ф и г. 36. Двухзонная модель.
а — $T=300$ К; б — $T=75$ К; в — $T=1200$ К.

намного выше зоны акцепторных поверхностных состояний, то $Q_{ss} \rightarrow -e\bar{N}E_b$. Эта величина представляет собой полный заряд заполненных акцепторов (eN_A). Если принять, что ими являются *собственные* поверхностные состояния, то $N_A \sim 4 \cdot 10^{14}$ см⁻². Отсюда можно получить

эффективную массу поверхностных состояний и ширину зоны, так как по определению

$$E_b = \frac{N_A}{\bar{N}} \approx \frac{1,0 \cdot m_0}{m_s^*}. \quad (12.9)$$

При отношении $m_s^*/m_0 \sim 10$ ширина зоны поверхностных состояний оказывается равной $\sim 0,1$ эВ. Это, по-видимому, типично для Si и Ge. Напротив, отношение масс $\sim 0,2$ соответствует ширине зоны поверхностных состояний ~ 5 эВ, что более характерно для соединений типа AlBV_6 , как это будет показано впоследствии. Аналогично можно также показать, что при $\Delta E_F \rightarrow -\infty$ имеем $Q_{ss} \rightarrow +e\bar{N}E_b$, и полностью заряженными (положительно) оказываются донорные состояния. Для значений ΔE_F , лежащих внутри зоны поверхностных состояний, вместо (12.8) получаем

$$Q_{ss} \approx e\bar{N} \left(\Delta E_F - \frac{1}{2} E_{sg} \right).$$

Другими словами, Q_{ss} зависит почти линейно от ΔE_F в этом вырожденном интервале. Внутри энергетической щели между поверхностными зонами функция (12.8) ведет себя, как и (12.4), подобно гиперболическому синусу. На границе зоны

$$Q_{ss} = -e\bar{N}kT \ln 2.$$

Из фигуры видно, что даже если плотность состояний $\bar{N}$ постоянна, Q_{ss} меняется с T *внутри* зоны поверхностных состояний. Этот эффект возникает ввиду необходимости учета максвелловских «хвостов». Предельным случаем при $E_b \rightarrow 0$ будет, конечно, уже рассмотренное распределение *двух дискретных уровней*.

В качестве третьего простого примера рассмотрим случай, когда пара акцепторных и донорных поверхностных состояний образует слившуюся зону поверхностных состояний, как это показано на фиг. 34, в. Нейтральный уровень E_0 отвечает центру зоны. Если, как и раньше, $\bar{N}$ — плотность состояний на единицу энергии, а E_b — ширина зоны (на фигуре $E_b = 0,25$ эВ), то

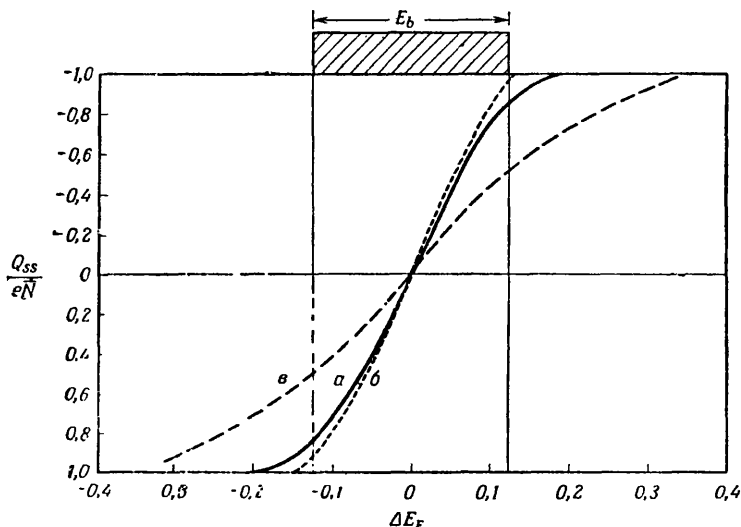
тогда получаем

$$Q_{ss} = -e\bar{N} \int_{-1/2 E_b}^{1/2 E_b} \left(1 + \exp \frac{E - \Delta E_F}{kT}\right)^{-1} dE + \\ + e\bar{N} \int_{-1/2 E_b}^{1/2 E_b} \left(1 + \exp \frac{\Delta E_F - E}{kT}\right)^{-1} dE. \quad (12.10)$$

Интегрирование дает

$$\frac{Q_{ss}}{e\bar{N}E_b} = \frac{2kT}{E_b} \ln \frac{\operatorname{ch} \left[\left(\frac{1}{2} E_b - \Delta E_F \right) / 2kT \right]}{\operatorname{ch} \left[\left(\frac{1}{2} E_b + \Delta E_F \right) / 2kT \right]}. \quad (12.11)$$

Эта зависимость показана на фиг. 37 для тех же трех температур.



Фиг. 37. Модель слившейся зоны.

а — $T=300$ К; б — $T=75$ К; в — $T=1200$ К.

Сравнение фиг. 35—37 позволяет лучше всего выявить характерные черты различных распределений. При температуре 1200 К все три схемы дают величину Q_{ss} ,

по существу линейно зависящую от ΔE_F . Это обусловлено тем, что величина щели (а также ширина зоны) составляет примерно $2,5 kT$. Чтобы выявить различие между этими распределениями, необходимо проводить эксперименты при гораздо более низких температурах (для подавления «теплового шума»). Эти температуры легко оценить. Пусть, например, $E_{sg} = 0,06$ эВ, тогда $E_{sg} = 2,5 kT$ при комнатной температуре. Этому, вероятно, отвечает случай Ge, для которого невозможно однозначно определить распределение поверхностных состояний, до тех пор пока не проведены эксперименты при температурах значительно ниже комнатной.

Полезно рассмотреть формально самое общее распределение поверхностных состояний. Пусть при этом $\bar{N}_A(E)$ и $\bar{N}_D(E)$ — акцепторная и донорная плотности состояний на единичный интервал энергий соответственно. Энергия E изменяется от $-\infty$ до $+\infty$. На деле функции $\bar{N}_A(E)$ и $\bar{N}_D(E)$ будут отличны от нуля только в ограниченной области вблизи объемной запрещенной зоны. Имеем

$$\frac{Q_{ss}}{e} = - \int_{-\infty}^{\infty} f(E - E_F, T) \bar{N}_A(E) dE + \\ + \int_{-\infty}^{\infty} f(E_F - E, T) \bar{N}_D(E) dE, \quad (12.12)$$

где $f(X, T)$ — функция распределения Ферми

$$f(X, T) \equiv \left[1 + \exp\left(\frac{X}{kT}\right) \right]^{-1}.$$

Доноры могут быть связаны с акцепторами, поскольку

$$f(X, T) = 1 - f(-X, T),$$

так что вместо (12.12) получаем

$$\frac{Q_{ss}}{e} = - \int_{-\infty}^{\infty} f(E - E_F, T) [\bar{N}_A(E) + \bar{N}_D(E)] dE + N_D, \quad (12.13)$$

где

$$N_D = \int_{-\infty}^{\infty} \bar{N}_D(E) dE. \quad (12.14)$$

Удобнее всего проанализировать выражение (12.13), разбивая функцию Ферми $f(X, T)$ на две части

$$f(X, T) \equiv f_0(X, 0) + f_1(X, T), \quad (12.15)$$

где f_0 — ступенчатая функция

$$f_0(X, 0) = \begin{cases} 1, & X < 0 \\ 0, & X > 0. \end{cases} \quad (12.16)$$

Как f_0 , так и f_1 имеют разрывы в точке $X = 0$. Объединяя (12.13) и (12.15), получаем

$$\begin{aligned} \frac{Q_{ss}}{e} = & - \int_{-\infty}^{E_F} [\bar{N}_A(E) + \bar{N}_D(E)] dE - \\ & - \int_{-\infty}^{\infty} f_1(E - E_F, T) [\bar{N}_A(E) + \bar{N}_D(E)] dE + N_D. \end{aligned} \quad (12.17)$$

Продифференцировав это выражение, имеем

$$\begin{aligned} \frac{\partial Q_{ss}/e}{\partial E_F} = & - [\bar{N}_A(E_F) + \bar{N}_D(E_F)] - \\ & - \int_{-\infty}^{\infty} \frac{\partial f_1(E - E_F, T)}{\partial E_F} [\bar{N}_A(E) + \bar{N}_D(E)] dE. \end{aligned} \quad (12.18)$$

Первый член в правой части (12.8) определяет вклад в заряд суммы донорных и акцепторных состояний, действительно находящихся в точке E_F . Второй член отвечает вкладу максвелловских «хвостов», охватывающих состояния, расположенные на некотором удалении от E_F . При $T = 0$ этот второй член должен исчезнуть. Из этого анализа следует, что было бы ошибкой находить величину $\bar{N}_A(E_F) + \bar{N}_D(E_F)$ простым дифференцированием Q_{ss} , как это могло бы показаться на первый взгляд. Второе слагаемое в (12.18) зачастую оказывается существенным, особенно при малых значениях величин E_{sg}/kT или E_b/kT . Это достаточно ясно показано на фиг. 35—37.

Хотя величина $\bar{N}(E)$ обычно предполагается не зависящей от температуры, заряд $Q_{ss}(E_F, T)$ зависит от температуры. Поэтому все эксперименты, которые интерпретировались в рамках модели с постоянным значением $\bar{N}(E)$ в узком энергетическом интервале, могут быть повторены при более низкой температуре для определения любой структуры, содержащейся в $\bar{N}(E)$. Важно отметить, что распределение $\bar{N}(E)$ можно экспериментально исследовать с помощью «спектрометров» с «разрешением» в несколько kT . Отметим также, что фиг. 36 и 37 показывают наличие зависимости от температуры даже в случае вырожденных поверхностных состояний (когда E_F лежит внутри зоны поверхностных состояний).

§ 13. Экспериментальные методы

Теперь мы кратко остановимся на различных методах, используемых для определения энергетического распределения поверхностных состояний. Экспериментальная техника, применяемая для изучения *несобственных* поверхностных и *контактных* состояний, интересна сама по себе, но она не пригодна для исследования собственных поверхностных состояний, которые мы здесь рассматриваем. Все методы, обсуждаемые ниже, делятся на четыре группы: легирование, наложение внешнего поля, облучение светом и все прочие. Опишем теперь эти методы в самых общих чертах, не вдаваясь в некоторые тонкости и детали, конкретное рассмотрение которых мы отложим до § 17.

1. Легирование

Весьма эффективным представляется метод, который состоит в выполнении одинаковых экспериментов на кристаллах с различной степенью легирования, так что покрывается интервал от вырожденного n -типа до вырожденного p -типа. Степень легирования связана с величиной ξ , согласно (11.2). Иногда используют кристаллы с неоднородным распределением примеси, таким, например, как в $p-n$ -переходе. Результаты измерений для

разных областей такого кристалла отвечают разной степени легирования. Иногда экспериментально бывает невозможно менять концентрацию примеси в заметной степени. Это имеет место в соединениях элементов II—VI и в некоторых случаях III—V групп. Например, CdSe может быть только *n*-типа, а ZnTe — только *p*-типа. Это положение, может быть, удастся изменить в будущем с помощью метода «пересадки» ионов [183].

Величина V_D может быть получена непосредственно для любой степени легирования путем измерения ϕ_w и Φ с использованием уравнения (11.6). Затем, согласно (11.5), можно вычислить поверхностный потенциал ϕ_s как функцию ζ . Поверхностный потенциал определяет положение уровня Ферми на поверхности по отношению к центру запрещенной зоны (точнее, по отношению к E_i). Зная ϕ_s , можно теоретически восстановить плотность объемного заряда Q_{sc}/e (см⁻²). Так как внешнее поле отсутствует, то $Q_{sc} = -Q_{ss}$; таким образом, заряд поверхностных состояний Q_{ss} можно представить как функцию ϕ_s ($\phi_s = E_F - E_i$) для различных значений ζ . График зависимости Q_{ss} от ϕ_s часто напоминает кривые фиг. 35 и 36, некоторые имеют сходство с гиперболическим синусом. Подбирая соответствующую кривую, можно найти распределение поверхностных состояний [175—177].

Менее удовлетворительные результаты получаются при измерении малых величин. Измерение фотоэмиссии (вместо ϕ_w) как функции легирования [184—186] дает некоторую информацию об искривлении зон, если оно велико и если зоны загнуты у поверхности вниз. Это обусловлено тем, что объемные электроны, проходя в слой объемного заряда, влияют на величину и форму кривой фотоэмиссии Y как функции $h\nu$. Может случиться, что измеряется именно ϕ_w (а не Φ) [137] как функция ζ . Для кристаллов с переменной концентрацией примеси — от вырожденного *n*-типа до вырожденного *p*-типа — изменение работы выхода $\Delta\phi_w$ в отсутствие поверхностных состояний будет, вероятно, равно E_g . Если же окажется, что $\Delta\phi_w \ll E_g$, то можно сказать, что зоны «закреплены» у поверхности поверхностными состояниями. Обратной мерой степени закрепления слу-

жит отношение $\Delta\phi_w/E_g$. Позже будет показано, что это отношение составляет $\sim 6\%$ в случае Ge и $\sim 20\%$ для Si.

Величину V_D также можно определить как функцию концентрации, не обращаясь к измерению ϕ_w или Φ . Можно, например, получить V_D из изменения проводимости при скалывании [178, 187], если принять подходящие модели рассеяния на поверхности. Таким образом, проходя уровнем Ферми через поверхностные состояния, можно детектировать распределение поверхностных состояний, как это видно из фиг. 35—37.

2. Внешние поля

Вместо того чтобы менять концентрацию примеси, можно варьировать внешнее поле, приложенное к образцу данной степени легирования. Обычно для этого используют электрод (как правило мегаллический), расположенный параллельно плоскости поверхности, как это показано на фиг. 33.

Рассмотрим сначала измерения эффекта поля при постоянном токе. Соответствующее уравнение (теорема Гаусса), связывающее Q_{sc} и внешнее поле E в отсутствие поверхностных состояний, имеет вид

$$Q_{sc} = K\varepsilon_0 E. \quad (13.1)$$

Для диэлектрической проницаемости $K = 10$ и напряженности поля $E = 10^5$ В/см, $Q_{sc} = 8,85 \cdot 10^{-8}$ Кл/см². Это отвечает плотности зарядов $Q_{sc}/e = 5,5 \cdot 10^{11}$ см⁻². Практически 10^5 В/см — это максимально достижимое значение напряженности поля в случае чистых поверхностей (без твердых прослоек, таких, как майлар, слюда или слой окисла). Оно соответствует, скажем, падению напряжения 10^3 В в вакуумном промежутке шириной 10^{-2} см. Таким образом, если имеются собственные поверхностные состояния, которые могут дать значительно большую плотность, чем $5 \cdot 10^{11}$ см⁻², эффект поля с постоянным током не может воздействовать или адекватно реагировать на эти состояния. По этой причине следует говорить, что число поверхностных состояний по крайней мере больше, чем $5 \cdot 10^{11}$ см⁻² (0,1% монослоя). Для

сравнения укажем, что изменение степени легирования в пределах вырождения позволяет зондировать плотности вплоть до 10% монослоя [175—177]. [Ситуация существенно меняется, если мы имеем дело с *реальными* (обычно загрязненными) поверхностями с прослойками майлара или слюды. Наибольшее поле здесь $\sim 2 \cdot 10^6$ В/см и $Q_{ss}/e = 10^{13}$ см $^{-2}$ (2,5% монослоя). Такое поле близко к разрушающему полю для большинства твердых тел.] В случае собственных поверхностных состояний в Si, Ge или соединениях типа $A^{III}B^V$ и $A^{II}B^{VI}$ может оказаться, что часть поверхностных состояний расположена вне запрещенной зоны. Эту часть распределения нельзя исследовать непосредственно ни с помощью эффекта поля, ни изменением степени легирования.

Ввиду невозможности применить эффект поля с постоянным током для эффективного зондирования поверхностных состояний в Si или Ge очень часто используется метод возмущения. Этот метод заключается в нахождении производной изменения поверхностной проводимости $\Delta\sigma$ по изменению заряда Q_f , накопленного на электроде. «Подвижность в эффекте поля» определяется тогда выражением

$$\mu_{fe} \equiv - \frac{d\Delta\sigma}{dQ_f}. \quad (13.2)$$

Это не есть истинная подвижность, так как не весь заряд Q_f подвижен. Как отмечалось ранее, заряды в поверхностных состояниях Q_{ss} считаются неподвижными. Для случая когда поверхностные состояния доминируют, а поверхность сильно насыщена, но не вырождена, можно показать, что¹⁾

$$\frac{dQ_{ss}}{du_s} \sim \frac{1}{2} \frac{\Delta\sigma}{\mu_{fe}}, \quad (13.3)$$

где u_s — нормированный потенциал поверхности ($u_s = \phi_s/kT$). Удобно интерпретировать результаты в предположении о постоянной плотности состояний $\bar{N}$ в интервале нескольких kT около E_0 , как показано на

¹⁾ См. работу [7], стр. 427.

фиг. 37. Тогда

$$\frac{dQ_{ss}}{du_s} = e\bar{N}kT,$$

и вместо (13.3) в модели слившейся зоны получаем

$$\bar{N} = -\frac{1/2\Delta\sigma}{ekT\mu_{fe}}. \quad (13.4)$$

Для Si [188] $\Delta\sigma = 0,59$ мкСм/ед. площади, $\mu_{fe} = -1$ см²/(В·с) и $kT = 0,025$ эВ, откуда $\bar{N} = 1 \cdot 10^{14}$ см⁻²·эВ⁻¹. Если эта плотность обусловлена слившейся зоной собственных поверхностных состояний, как на фиг. 37, то эффективная масса, в соответствии с (12.7) должна определяться выражением $m_s^*/m_0 \sim 0,24$, что отвечает, согласно (12.9), ширине зоны $E_b \sim 5$ эВ.

С другой стороны, выберем модель двух дискретных уровней, как на фиг. 35, и пусть уровень Ферми находится около центра запрещенной зоны между поверхностными состояниями, $\Delta E_F = 0$. Теперь, пользуясь формулой (12.5) вместе с (13.3), можно получить иную трактовку тех же данных по эффекту поля:

$$N = -\left(\frac{1/2\Delta\sigma}{e\mu_{fe}}\right)\left[1 + \operatorname{ch}\left(\frac{1/2E_{sg}}{kT}\right)\right]. \quad (13.5)$$

Если $E_{sg} = 0,25$ эВ и $kT = 0,025$ эВ, то множитель в скобках равен 75 и $N = 2 \cdot 10^{14}$ см⁻², что можно ожидать в случае примерно одного поверхностного состояния на поверхностный атом. Как показывают другие эксперименты с Si, модель *двух дискретных уровней*, видимо, более правильна, чем модель *слившейся зоны*. Проведенный расчет говорит о том, что измерение подвижности в эффекте поля имеет ограниченную применимость для нахождения распределений поверхностных состояний. Это связано с тем, что измерение проводится только в одной точке. Однако рассмотрение фиг. 35—37 показывает, что дополнительные измерения температурной или концентрационной зависимости могут помочь различить детали структуры поверхностных состояний. В связи с интерпретацией результатов по подвижности в эффекте поля используется и другая модель —

одионочная дискретная зона с центром в точке E_F , как показано на фиг. 37, и с шириной $E_b = 0$. При этом число поверхностных состояний равно¹⁾ $N = 4kT\bar{N}$, где $\bar{N}$ — плотность состояний на единицу энергии, взятая из (13.4). Франкль [8, 189, 190]²⁾ весьма подробно рассмотрел критерии оценки подвижности в эффекте поля. Он нашел, что в некоторых случаях, данные по которым были сообщены и проанализированы другими авторами, результаты не являются согласованными и противоречат свойствам функции Ферми. Именно всегда должны выполняться соотношения

$$f \geq \frac{df}{du_s} \quad \text{и} \quad \frac{df}{du_s} \leq \frac{1}{4}.$$

Измерения импульсного эффекта поля [191—193] были использованы для определения динамической реакции занятых и незанятых поверхностных состояний. Исследуя зависимость постоянной времени от $1/T$, можно получить значения эффективного энергетического уровня Аррениуса ($E_2 - E_V$ или $E_c - E_1$ на фиг. 34, а) и сечения захвата. Акцепторные (донорные) состояния обычно оказываются расположенными ближе к зоне проводимости (валентной зоне), как показано на фиг. 34, а. Исследования с эффектом синусоидального поля выполнялись для различных частот и помогли разделить состояния на быстрые и медленные. В настоящее время установлено [194], что на чистых поверхностях нет медленных состояний. Такие состояния появляются только на окисленных или загрязненных поверхностях полупроводников. Их называют «медленными состояниями» ввиду их слабого взаимодействия с контактом.

Если постоянный ток в эффекте поля достаточно велик, может возникнуть холодная эмиссия. При этом иногда удается обнаружить поверхностные состояния. Такие эксперименты по холодной эмиссии для Ge были проделаны Шепардом [195]. (См. также работу Артура [196].)

¹⁾ См. примечание на стр. 154.

²⁾ См. также [8], стр. 212, 213.

3. Облучение

Облучение часто используется в качестве инструмента при исследовании поверхностных состояний. О некоторых эффектах, вызванных облучением, мы уже кратко говорили. Это касается эффекта фото-э.д.с. и фотоэмиссии в вакууме. В эффекте фото-э.д.с. измеряется только величина V_D . При этом непосредственно измерить величину взаимодействия света с поверхностными состояниями не удастся. При фотоэмиссии возможно в принципе возбуждение электронов из заполненных поверхностных состояний в вакуумную область. Если заполненные поверхностные состояния расположены выше валентной зоны, выход фотоэлектронов Y должен иметь «хвост» в области низких энергий. Этот «хвост» может быть обнаружен на опыте. Долю выхода, обусловленную поверхностными состояниями, в настоящее время не удастся оценить вследствие незнания сечения реакции. Одна из главных проблем, стоящих при использовании метода фотоэмиссии, — выделить вклад поверхностных состояний на фоне вклада объемных. Измеряемый квантовый выход Y (отношение числа эмиттированных электронов к числу поглощенных) часто имеет вид [197]

$$Y = C_1 (h\nu - \Phi)^{r_1} + C_2 (h\nu - \Phi_2)^{r_2}, \quad (13.6)$$

где Φ и Φ_2 — соответственно нижнее и верхнее значения энергий, а r_1 и r_2 — целые или полуцелые числа (1, $3/2$, 2, $5/2$, 3 и т. д.). Иногда, как, например, для Ge, в формуле (13.6) появляются три слагаемых. В других случаях [198, 199] считают, что эти слагаемые входят в формулу не аддитивно, а переходят одно в другое, так что и сам выход Y и его первая производная остаются непрерывными в некоторой точке $h\nu_0$. Исходя из различных механизмов, были проведены теоретические оценки [200—202] показателей r_1 и r_2 . Прямые переходы в объеме соответствуют значению $r_2 = 1$. Иногда, например для очищенного ионной бомбардировкой и отожженного Si, линейный участок отсутствует [203]. Низкоэнергетический хвост может иметь показатель $r = 3$, который трудно отличить на опыте от значения $r = 5/2$. Участок, отвечающий этому «кубическому закону» или

«закону $^{5/2}$ », по-видимому, наблюдается у всех веществ [154, 204, 205]. Кейн [200] теоретически вычислил показатели степени, которые можно было бы ожидать для одиннадцати различных механизмов процесса. Ни один из них не дает $r = 3$. Рассмотренные им механизмы включают различные комбинации переходов в объеме и на поверхности, прямые и не прямые переходы, переходы с рассеянием и без него. Тем не менее большинство исследователей в настоящее время принимают кубический закон ($r = 3$) как удобную форму для параметризации полученных фотоэмиссионных данных вблизи порога. Здесь мы будем придерживаться той же линии. Для теоретического расчета выхода фотоэмиссии требуется знание матричного элемента между начальным и конечным состояниями, а также распределение всех этих состояний по энергии [201, 202]. Таким образом, вид формулы (13.6) и соответствующая величина r зависят от зонного спектра кристалла в области энергий вблизи значения E' (для вакуума), как на фиг. 33. Фишер [206] дал обзор фотоэмиссионных процессов и сделал вывод, что фотоэмиссия с поверхностных состояний может быть в принципе выделена только как отклонение от закона кубического корня для хвоста фотоэмиссии. Для этого требуются чистые и не аморфные образцы [207]. Энергетическое распределение эмиттированных фотоэлектронов было получено с использованием задерживающего потенциала [203]. Анализ этого распределения оказывает существенную помощь в интерпретации механизмов эмиссии. Однако до настоящего времени эмиссия с поверхностных состояний не была однозначно идентифицирована [154, 204, 206].

Поверхностные состояния могут также взаимодействовать с излучением, квант энергии которого меньше ширины запрещенной зоны, $h\nu \leq E_g$. Так, может иметь место фотоионизация заполненных поверхностных акцепторов ($h\nu > E_c - E_1$ на фиг. 34, а) или фотоионизация поверхностных доноров ($h\nu \geq E_2 - E_v$ на фиг. 34, а). Эти и подобные им переходы могут быть измерены спектроскопическими методами в условиях полного внутреннего отражения [208—210], методами фотоиспарения [211] и фотодесорбции [212, 213].

4. Другие методы

Для исследования распределения поверхностных состояний может использоваться изменение температуры. Этот метод основан на температурной зависимости параметра легирования ζ и заряда Q_{ss} , как это следует из фиг. 35—37. Заполненные поверхностные ловушки могут быть ионизированы нагревом [214] и «термически индуцированный ток» можно измерить в фотопроводниках [215].

Сведения о заселенности поверхностных состояний были получены при измерениях электронного спинового резонанса на свежесколотых, раздробленных или отожженных кристаллах [216—219]. Надлежащим образом калиброванный эталон позволяет определить число спинов на 1 см^2 , которое в свою очередь связано с числом *неспаренных* поверхностных состояний.

По аналогии с примесными зонами, которые, как известно, образуются в сильно легированных полупроводниках, возможна также проводимость за счет зон поверхностных состояний [178, 179, 187, 220]. Этот эффект исследовался в случае Si и Ge, однако проводимости у поверхностных состояний найдено не было. Причины этого еще не ясны. Для исследования поверхностных состояний также применялся эффект туннелирования [221, 222] в $n-p-n$ - и $p-n-p$ -контактах.

В присутствии *некоторых типов* поверхностных состояний резко усиливается рекомбинация избыточных объемных носителей (созданных, например, облучением). Поверхностные состояния действуют здесь как центры рекомбинации. Вообще говоря, эффективный центр поверхностной рекомбинации должен захватывать и дырки, и электроны с заметным сечением в каждом случае. Изгиб зоны и подвижности в объеме также должны быть такими, чтобы избыточные электроны и дырки с легкостью могли достигать поверхности. Следует, однако, отметить, что не все поверхностные состояния оказываются хорошими центрами рекомбинации. Таким образом, рекомбинационный метод позволяет исследовать только часть распределения поверхностных состояний.

§ 14. Приготовление поверхности

Вряд ли есть какой-нибудь смысл проводить точные эксперименты, если объектом исследования являются поверхности, строение которых плохо известно. Собственные поверхностные состояния легко могут быть замаскированы несобственными поверхностными состояниями, поверхностными окислами, металлами, адсорбированными из травителей, и т. д. Поэтому приходится уделять особое внимание приготовлению чистых поверхностей.

Хорошим способом приготовления чистой поверхности является скол кристалла в условиях сверхвысокого вакуума. Скалывание легко осуществить только по некоторым плоскостям — специфическим для каждого типа кристаллов. Вот наиболее важные примеры:

- (111) кристаллы IV группы (типа алмаза), Si, Ge,
- (110) соединения типа $A^{III}B^V$ и $A^{II}B^{VI}$ (цинковая обманка), GaAs, CdTe,
- ($11\bar{2}0$) соединения типа $A^{II}B^{VI}$ (вюрцит), CdS, CdSe,
- (100) соединения типа $A^{IV}B^{VI}$ и $A^I B^{VII}$ (каменная соль), PbS, AgBr.

Иногда встречаются важные исключения. Например, ZnO (структура вюрцита) в игольчатой форме имеет призматические грани типа $\{10\bar{1}0\}$ и скалывается поперек иглы вдоль плоскости $\{0001\}$ (плоскость скола перпендикулярна оси c) [223]. Степень совершенства скола зависит от геометрии. Как показывают данные, полученные с помощью оптической и электронной микроскопии [203], а также зависимость Y от $h\nu$ [203], некоторые типы скола получаются лучше, чем другие. Для получения хорошего скола может потребоваться особая геометрия кристалла и резца. Плохо сколотые (или сложенные) поверхности иногда дают кривые $Y(h\nu)$, выгнутые вверх [184—186, 204]. Поверхности готовят также путем скола в относительно инертном жидком N_2 [224] или в органических жидкостях [225]. Чистота этих поверхностей несколько сомнительна, так что возможно присутствие несобственных поверхностных состояний. Совершенно ясно, что поверхности, пригото-

ленные на воздухе и не очищенные вновь, почти бесполезны при изучении собственных поверхностных состояний.

Другим способом получения достаточно чистой поверхности (свободной от посторонних атомов) является «бомбардировка и отжиг». При этом вырезается кристалл с требуемой поверхностью, которую обтачивают, полируют, протравливают и помещают в сверхвысокий вакуум. После этого поверхность подвергают ионной бомбардировке (обычно ионами аргона) и затем отжигают. Этот цикл повторяется несколько раз. Впервые такая процедура была предложена Фарнсворсом с соавторами [226, 227]. Метод бомбардировки и отжига устраняет повреждения при обработке, окисление, загрязнение травителем и т. п. Обычно в результате получается достаточно чистая поверхность, что подтверждается результатами экспериментов по ДЭНЭ.

Нагрев в сверхвысоком вакууме кристаллов, выдержанных в воздушной среде, *редко* дает атомно чистую поверхность. Фактически на поверхности к исходному загрязнению, обусловленному выдержкой на воздухе, лишь *добавляется* загрязнение, источники которого следующие: 1) дегазация и диффузия примесей из объема к поверхности и 2) выделение газов со стенок камеры, находящихся вблизи образца. Загрязненная таким образом поверхность может быть очищена чередованием бомбардировки ионами с отжигом. Но вряд ли одного нагрева будет достаточно.

Итак, можно сделать заключение, что только два метода заслуживают определенного доверия как методы получения собственных поверхностей: метод скола и метод бомбардировки и отжига. (Эти сами по себе сложные и требующие значительного времени методы сталкиваются еще и со своими специфическими проблемами. Возникают, например, трудности, связанные с получением идеального скола и возможностью образования после скола пленки p -типа на поверхности. Все это мы обсудим в § 17.) Около 5% работ по поверхностным состояниям используют указанные два метода, и, таким образом, речь в них идет именно о *собственных* поверхностных состояниях (данным работам посвящен

§ 17). В остальных 95% работ, в которых преследуются другие цели, эти два метода подготовки поверхности не используются. Соответствующие обширные обзоры опубликованы в [7, 8].

Сравнительно недавно появился полезный метод определения низких концентраций поверхностных примесей (менее 5% монослоя). Речь идет о спектроскопии Оже [228]. С помощью холодной эмиссии ионов [165, 166] можно получить сведения об атомной геометрии поверхности, включая расположение примесей. Однако этот метод не применяется для полупроводников из-за их недостаточной прочности в необходимых сильных электрических полях. Обзор многочисленных методов анализа поверхности сделан Марком и Левином в работе [229].

§ 15. Метод дифракции электронов низких энергий (ДЭНЭ)

Мощным инструментом, предназначенным для исследования структуры поверхностных атомов, является метод ДЭНЭ. Этот метод в гораздо большей степени пригоден для изучения свойств поверхностных состояний, чем, скажем, оптические методы. С помощью ДЭНЭ в CdS [230] наблюдалась перестройка поверхности и микрофасеточная структура размером менее 100 Å, что недоступно для лучших оптических микроскопов.

В методе ДЭНЭ коллимированный моноэнергетический пучок электронов падает на поверхность. Упруго рассеянные электроны регистрируются затем флуоресцирующим экраном или попадают в клетку Фарадея. Пучок проникает вглубь только на несколько атомных слоев, так что он «видит» не столько объем, сколько поверхность. Имеется обзор, посвященный методу ДЭНЭ и полученным с его помощью результатам [232]. С помощью ДЭНЭ в большинстве работ получают информацию в основном трех типов. Большинство дифракционных картин, подобно лауэграммам, представляет собой некоторую совокупность рефлексов. В этом случае, во-первых, положения рефлексов указывают на периодичность расположения атомов в решетке вдоль поверх-

ности. Поверхность, обладающая периодами, равными объемному периоду, дает на дифракционной картине известную систему рефлексов. Если же поверхность имеет, скажем, удвоенный атомный период, то получится дополнительная система рефлексов, расположенных между предыдущими. Во-вторых, отношения интенсивностей рефлексов указывают относительные доли различных структур. В-третьих, изменение интенсивности отдельного рефлекса как функция энергии падающего электрона может с помощью закона Брэгга многое сказать о самой структуре. К сожалению, метод ДЭНЭ встречается с трудностями, которых не знает дифракция рентгеновских лучей. Так, взаимодействие электронов друг с другом вызывает множественное рассеяние. Были развиты самосогласованные теоретические схемы [233], которые дали брэгговские пики дробного порядка, т. е. $n = 1, 1\frac{1}{3}, 2, 2\frac{2}{3}, \dots$, и т. д. Кроме того, атомный фактор рассеяния и поверхностный потенциал можно лишь оценить. Эта проблема стоит особенно остро при интерпретации данных по ДЭНЭ на поверхностях бинарных кристаллов.

Согласно результатам ДЭНЭ, все поверхности Si и Ge, по-видимому, испытывают перестройку [52, 149, 234, 235]. Другими словами, имеются четкие рефлексы, не отвечающие идеально оборванной решетке. В алмазе рефлексы перестройки очень слабы [53] (относительная интенсивность менее 1%). Оказывается также, что перестройка — обычное явление для «полярных» поверхностей, таких, как поверхности (111) и $(\bar{1}\bar{1}\bar{1})$ цинковой обманки [148, 149] и поверхности (0001) и $(000\bar{1})$ вюрцита [230, 231]. Все эти поверхности имеют только один сорт атомов в поверхностном слое и оказываются нестабильными. Это можно видеть уже исходя из простых электростатических соображений [52, 236]. Причины, вызывающие перестройку, в настоящее время еще не вполне поняты. В частности, сугубо приближенно можно принять, что смещение атомов возмущает ненасыщенные связи, и это ведет к их насыщению, или что сдвиг электрических зарядов приводит к более нейтральной поверхности. Детальные атомные схемы поверхностей

(111) Si, Ge, GaAs и GaSb были предложены Ландером и Моррисоном [52, 234, 235], Макрэем [149], Сиватцем [237] и Хейнеманом [146, 238]. Модели Ландера и Моррисона, а также Макрэя допускают, что некоторые атомы удалены. Модель Сиватца предполагает наличие сопряженных связей, подобно имеющимся у органических молекул. В модели Хейнемана поверхность представляется сморщенной или покоробленной, так что некоторые из *поверхностных* атомов поднимаются или опускаются на расстояние до 1 Å. Недавние работы по адсорбции Na [239], рекристаллизации после скола [240], изучению силы связей [241] и электронному спиновому резонансу [216—218] показывают, что в настоящее время в случае Si и Ge модель Хейнемана, по-видимому, является предпочтительной. Однако для интерпретации дифракционных картин, для какой бы поверхности данного кристалла (иногда и при данной температуре) они ни были получены, привлекаются спорные предположения. Поэтому мы здесь не будем пытаться однозначно связать атомную структуру с видом дифракционных картин.

«Ионные» или «скомпенсированные» поверхности частично ионных соединений, по-видимому, не перестроены [148, 242—244]. Новые рефлексy у них отсутствуют. Вероятно, это обусловлено электрической нейтральностью таких поверхностных плоскостей (равное число ионов A и B в поверхностной плоскости кристалла AB) [30, 236]. Из анализа интенсивностей, вероятно, следует, что, например, на поверхности (110) GaAs подрешетка Ga остается жесткой, но слегка сдвигается (менее чем на 1 Å) вдоль поверхности или по нормали к ней относительно подрешетки As [148, 149]. То же, по-видимому, наблюдалось у всех соединений типа $A^{III}B^V$ со структурой цинковой обманки [148, 243], а также для поверхности (10 $\bar{1}$ 0) CdS (структура вюрцита) [244]. (Эти кристаллические грани, если они идеально сколоты, должны, естественно, иметь правую и левую стороны¹⁾.) Прочие поверхности вовсе не возмущаются в продольном направлении. Одна из них — поверхность (100) NaCl

¹⁾ См. также стр. 128. — *Прим. перев.*

[242]. Существующее в этом случае растяжение решетки в поперечном направлении составляет несколько процентов [245, 246]. По сравнению с перестройкой или слабыми продольными сдвигами, описанными выше, это будет эффект второго порядка. Наконец, получаемые при сколе *перестроенные поверхности* имеют фазовые переходы при нагревании [52, 234, 235]. При этом меняется картина рефлексов. Каждой исследуемой поверхности может быть приписано обозначение, например Si (111)-7. Для неперестраиваемых поверхностей, по-видимому, нет фазовых переходов [148].

§ 16. Линейные дефекты

Слабым местом метода ДЭНЭ является его неадекватность при определении деталей поверхности, содержащих менее чем, скажем, 15% фактического числа поверхностных атомов [149]. Так, например, *линейные дефекты*, отвечающие краям ступенек или границам микрофасеточной структуры (сюда же относятся дислокации и границы зерен), если они затрагивают менее 1% поверхностных атомов, не проявляются в экспериментах по ДЭНЭ. Чтобы оценить вклад этих дефектов, примем, что тщательным образом сколотый по плоскости (111) кристалл Si имеет ступени шириной $2000 \pm 1000 \text{ \AA}$ и высотой более 50 Å. Эти оценки получены с помощью электронной микроскопии [203]. Число поверхностных атомов, приходящихся на линейные дефекты, будет равно

$$N_l \approx \left(\frac{1 \text{ линия}}{0,2 \cdot 10^{-4} \text{ см}} \right) \left(\frac{1 \frac{\text{число атомов}}{\text{линия}}}{3 \cdot 10^{-8} \text{ см}} \right) \approx \approx 1,6 \cdot 10^{12} \frac{\text{число атомов}}{\text{см}^2}. \quad (16.1)$$

Для сравнения укажем, что на нормальной поверхности имеется $8 \cdot 10^{14}$ атомов на 1 см^2 . Таким образом, краевые атомы составляют около 0,2% монослоя. Аналогичные данные получены для самых лучших сколотых кристаллов типа As_2V , где ступени имеют ширину $7000 \pm 3000 \text{ \AA}$ и высоту 50—200 Å [148]. Предполагается, что с краевыми атомами связаны поверхностные

состояния. Заряды этих состояний должны компенсироваться почти цилиндрической пространственно-заряженной оболочкой. Поверхность получается при этом как бы «пятнистой». Это пятнистое строение поверхности должно, по-видимому, проявляться в электрических и оптических измерениях. Полезно сопоставить эти ожидаемые «ступенчатые» и «краевые» состояния с состояниями у краевых дислокаций, которые, как известно, существуют в объеме Ge. Пирсон и др. [247, 248] обнаружили и объяснили акцепторные свойства состояний у краевых дислокаций в объеме. Шокли [249] предсказал, что эти состояния образуют примесную зону и ответственны за низкотемпературную объемную проводимость Ge. В противоположность им поверхностные состояния не дают вклада в проводимость [178, 220]¹⁾. Играс [250] изучал состояния у краевых дислокаций с помощью техники ямок травления и метода «электронного зеркала».

Итак, оказывается, что краевые атомы не влияют на дифракционную картину, полученную методом ДЭНЭ [52, 234, 235], или на зависимость Y от $h\nu$ для таких хорошо сколотых кристаллов [175, 176, 204]. Кроме того, заряды, измеренные в поверхностных состояниях Si, соответствуют более 10% монослоя [176] ($Q_{ss}/e > 0,4 \cdot 10^{14}$ число состояний/см²). Это более чем в 50 раз превышает плотность краевых атомов. Сванк также обнаружил сходство результатов по измерению Φ и ϕ_w для кристаллов типа $A^{IV}B^{VI}$ с различным качеством хорошо сколотых поверхностей [154].

§ 17. Результаты и их интерпретация

Как видно из предыдущих параграфов, получение данных по собственным поверхностным состояниям фактически довольно затруднительно с экспериментальной точки зрения, причем возникают определенные сложности в отношении их анализа и интерпретации. Каждый из экспериментаторов сообщает свои собственные исходные данные, свою интерпретацию и свою собственную

¹⁾ Исследование сколотого кристалла ZnO при 90 К было проведено в работе [223].

процедуру получения чистой поверхности. Все это имеет свои за и против. Например, сколотые поверхности дают более воспроизводимые свойства фотоэмиссии, чем поверхности, полученные бомбардировкой и отжигом [203]. Однако первые иногда дают менее хорошо воспроизводимую картину ДЭНЭ, чем последние [251]. Поэтому ясно, что любой эксперимент, связанный с распределением поверхностных состояний, можно подвергнуть критике с различных точек зрения. Иногда эта критика может быть не очень глубокой (как, например, в отношении энергетического сдвига всего распределения); в других случаях она оказывается серьезно обоснованной (например, при использовании слюдяной прослойки, находящейся в близком контакте с поверхностью). Поэтому было бы нелогично воспроизводить здесь все полученные результаты даже для чистых поверхностей.

Вместо этого здесь приведены только те эксперименты, *обработка которых требует привлечения теоретических моделей собственных поверхностных состояний*. Другие экспериментальные работы, касающиеся адсорбции, контактных состояний, а также физических, химических и структурных аспектов поверхностей, будут намеренно опущены, чтобы не загромождать картину. Мы остановимся в особенности на электрических свойствах свободных собственных поверхностей.

Примем для простоты, что нуль энергии находится вблизи центра запрещенной зоны $E_i = 0$. В частности, возможны следующие положения уровня Ферми у поверхности: $E_F = 0$ (собственная поверхность), $E_F > 0$ (поверхность n -типа), $E_F < 0$ (поверхность p -типа). Мы приведем по возможности результаты всех экспериментаторов, не вдаваясь, однако, в детальное рассмотрение способов их получения. Это даст возможность сэкономить немало места и времени и позволит молодому исследователю охватить всю область в целом. В этом параграфе будут рассмотрены кремний, германий и соединения элементов III—V и II—VI групп в указанном порядке. В заключение с целью сравнения будет добавлен небольшой раздел, касающийся несобственных поверхностных состояний.

1. Кремний

Распределение собственных поверхностных состояний в Si известно, вероятно, лучше, чем во всех других веществах. В основном оно было определено в 1962 г. Алленом и Гобели [175—177]. За этими первыми работами последовало много других, но результат существенно не изменился. Напротив, многие эксперименты по изучению электрических и оптических свойств, а также по электронному парамагнитному резонансу в целом подтвердили найденное распределение или по крайней мере не очень сильно ему противоречили.

Изучение кремния осложняется тем, что его поверхность перестроена. После скола в сверхвысоком вакууме при комнатной температуре дифракционная картина, полученная методом ДЭНЭ, характеризуется обозначением Si (111)- $1 \times \sqrt{3}$ или Si (111)- 1×2 в зависимости от того, выбраны ли в качестве основных наклонные оси [216] или нет [52, 234, 235]. После скола и нагрева эта картина необратимо переходит в такую, которая отвечает структуре Si (111)-7, т. е. очевидно, в ту же структуру, которая получается после ионной бомбардировки и отжига. Дополнительные сведения, относящиеся к этим структурам, а также к промежуточной структуре Si (111)-5, можно найти в работах Ландера [52] и Ландера и Моррисона [234]. Скалывание кремния тоже довольно трудная операция. При некоторых геометриях скола результаты *гораздо* лучше, чем при других [175—177, 203]. Теоретическое рассмотрение поверхностных состояний в Si также является непростой задачей, так как запрещенная зона в объеме *не прямая* ($E_g = 1,09$ эВ при комнатной температуре). По этим причинам Si, видимо, не является лучшим объектом для выполнения основных экспериментов по поверхностным состояниям. Имеются, однако, и положительные стороны. Их пять: практическая ценность, контролируемая чистота, контролируемые *n*- и *p*-примеси, отсутствие ловушек в объеме и простой вид связи в объеме (чисто ковалентная связь). С другой стороны, ионным плоскостям соединений элементов III—V и II—VI групп, имеющих прямые запрещенные связи, отвечают неперестро-

енные поверхности. Однако объемные свойства этих соединений, в том числе концентрация примесей, труднее поддаются контролю, чем в случае Si (или Ge). Повторим еще раз, что здесь нами рассматриваются только *собственные* поверхностные состояния. Обширную литературу, касающуюся адсорбционной системы Si — кислород и контактов Si — SiO₂, мы не будем рассматривать вплоть до п. 4 настоящего параграфа.

Рассмотрим качественные свойства сколотых поверхностей Si, следуя Аллену и Гобели [175—177], Ван-Лаару и Ширу [184—186], Хандлеру [187], Аспнесу и Хандлеру [178], Хенслеру [188], Хенслеру и Хейланду [252], Хейнеману [216] и Фишеру [253]. Оказывается, что поверхностные состояния распределены донорно-акцепторными парами, как предполагалось в *модели двух дискретных уровней* (фиг. 34, а), *модели двух зон* [$E_b \sim 0,2$ эВ] (фиг. 34, б) или в *sh-образной модели* (фиг. 34, г). При переходе от вырожденного *n*-типа к вырожденному *p*-типу уровень Ферми смещается только на величину $\sim 0,2$ эВ, отсюда $E_{sg} \sim 0,2$ эВ. Эти данные не позволяют в настоящее время сделать дальнейший выбор между указанными тремя или другими аналогичными моделями. Основной довод в пользу перечисленных моделей тот, что график зависимости Q_{ss} от E_F имеет вид sh-образной кривой [176], которая теоретически предсказана для этих моделей [см. (12.4), (12.8) и (12.17)]. А именно наклон кривой (dQ_{ss}/dE_F) в центре E_0 распределения поверхностных состояний минимален. Таким образом, модель слившихся зон (фиг. 35, в) может быть в случае сколотого Si отброшена [см. (12.11)]. Распределения поверхностных состояний для сколотого и нагретого Si или Si, приготовленного бомбардировкой и отжигом, значительно менее определены. Дело в том, что соответствующие эксперименты для различных концентраций примеси не проводились.

Возникает вопрос: являются ли наблюдаемые поверхностные состояния собственными или же они обусловлены поверхностными дефектами, такими, как ступеньки скола, тепловые ямки травления или краевые дислокации? Решить его можно двумя способами. Во-первых, полученное число поверхностных состояний было

$N \geq 10^{14} \text{ см}^{-2}$. Это во много раз больше, чем число поверхностных ступенек, ямок травления или дислокаций, на что уже указывалось в § 16. Во-вторых, наличие знака неравенства у большинства проанализированных данных не противоречит предположению о том, что число собственных поверхностных состояний $2N = 8 \cdot 10^{14} \text{ см}^{-2}$. Как видно из фиг. 35—37, число поверхностных состояний можно определить, только если удастся «пройти» уровнем Ферми *через* эти состояния. Поскольку в случае Si это сделано не было, экспериментаторам приходится ставить знак $\geq$. Проблема особенно актуальна при измерениях подвижности в эффекте поля, которые дали максимальное число поверхностных состояний в интервале порядка $\pm kT$ около уровня Ферми. С этой точки зрения важность измерений подвижности в эффекте поля [7, 8] несколько переоценивалась. Вместе с тем для постулированного распределения в сколотом Si измерения подвижности в эффекте поля, проводимые вблизи минимума плотности состояний ($E_F = E_0$), должны давать сильно заниженное число имеющихся поверхностных состояний. Такая ситуация характерна для кристалла Si *p*-типа с сопротивлением 200 Ом·см. Существует один случай, когда вводилось обратное неравенство $\leq$. Оно было предложено в порядке обсуждения Фишером [253] для объяснения температурной нечувствительности. Этот вопрос будет рассмотрен ниже более детально. Динамические свойства поверхностных состояний в Si довольно необычны. Фотоэмиссия возникает при наличии в этих состояниях электронов, но *только* для образцов *n*-типа [175, 176] (на сколотой поверхности). Эти состояния, по-видимому, будут не проводящего типа [178], несмотря на то, что Q_{ss}/e может достигать 10% монослоя в вырожденных образцах *n*-типа [175, 176]. Сигнал ЭПР после скола отвечает плотности состояний порядка 10^{14} см^{-2} , что указывает на собственные поверхностные состояния [216]. После нагрева в сверхвысоком вакууме амплитуда сигнала *уменьшается* на порядок или более [217, 218]. Интересно, что фотоэмиссия с поверхностных состояний у образцов *p*-типа *резко возрастает* после нагрева в сверхвысоком вакууме [203].

Итак, качественные характеристики поверхностных состояний в Si следующие: 1) распределение типа фиг. 34, *а, б* и *г*, но не *в*; 2) плотность достаточно высока, чтобы быть связанной с собственными поверхностными состояниями; 3) необычные динамические свойства. Положение уровня E_0 до некоторой степени спорно. Этот и другие вопросы разбираются ниже, где мы кратко рассмотрим отдельные работы.

Основным результатом Аллена и Гобели [175—177] был анализ функций ϕ_w и Φ для сколотого Si (кубический закон) в *непрерывном* интервале концентраций примеси от вырожденного *n*-типа до вырожденного *p*-типа. Это изменение концентраций способно сдвинуть уровень Ферми примерно на 0,2 эВ, что составляет лишь около 20% ширины запрещенной зоны. Таков результат «закрепляющего действия» поверхностных состояний. Для оценки распределения поверхностных состояний принималось, что энергия $\Phi = 5,15$ эВ — постоянная для всех концентраций и отсчитывается от края валентной зоны у поверхности. Эта величина Φ , вообще говоря, отличается от истинного значения. Наблюдаемое уменьшение истинной величины Φ в сильно легированных образцах *p*-типа обязано эмиссии с подповерхностных уровней (глубина выхода ~ 25 Å). Кроме того, оно вызывается уменьшением χ для сильно вырожденных образцов. Наблюдаемое уменьшение истинной величины Φ в сильно легированных образцах *n*-типа обусловлено фотоэмиссией с заполненных поверхностных состояний акцепторного типа. Соответствующая зависимость Y от $h\nu$ оказалась экспоненциальной, а не степенной, как предсказывалось Кейном [200]. (Все степенные зависимости имеют вертикальную асимптоту в полулогарифмическом масштабе [154].)

Энергия Φ фактически оказалась равной 5,15 эВ только для значений E_F в интервале $-0,4 \text{ эВ} < E_F < 0,0 \text{ эВ}$. Пользуясь значениями Φ и ϕ_w для различных концентраций, была вычислена зависимость Q_{ss} от E_F . График оказался sh-образного типа. Точнее говоря, данные были аппроксимированы кривой

$$\frac{Q_{ss}}{e} = -2,44 \cdot 10^{12} \text{ sh} \left[\left(\frac{E_F}{kT} \right) + 9,7 \right] \text{ см}^{-2},$$

причем E_0 расположено в точке $-9,7 kT = -0,23$ эВ. Поскольку результаты не слишком сильно отклоняются от sh-образной кривой в исследуемом интервале значений E_F , было сделано заключение, что состояния должны находиться вне рассмотренного интервала. Это согласуется с теоретически предсказанными зависимостями на фиг. 35 и 36. Для модели *двух дискретных уровней* значения параметров оказались следующими (см. фиг. 34, а и 35):

$$E_0 = -0,23 \text{ эВ}, \quad N_1 = N_2 \geq 1,8 \cdot 10^{14} \text{ см}^{-2}, \quad E_{sg} \leq 0,30 \text{ эВ}.$$

Для *двухзонной модели* (см. фиг. 34, б и 36) соответственно

$$E_0 = -0,23 \text{ эВ}, \quad \bar{N}_1 = \bar{N}_2 \geq 2,6 \cdot 10^{15} \text{ см}^{-2} \cdot \text{эВ}^{-1}, \\ E_{sg} = 0,2 \text{ эВ}, \quad E_b \leq 0,16 \text{ эВ}.$$

Если (в наихудшем случае) заменить неравенство для $\bar{N}_1$ равенством, то $N_1 = \bar{N}_1 E_b$ и

$$N_1 = N_2 = 4 \cdot 10^{14} \text{ см}^{-2}, \quad m_s^* = 6,5 m_0.$$

Большая эффективная масса означает большую плотность состояний на единицу энергии. Укажем для сравнения, что объемным состояниям отвечает малая эффективная масса и сравнительно большая ширина зоны, как это видно из (12.9). Третья из рассмотренных моделей характеризуется sh-образным распределением, как показано на фиг. 34, г. Минимуму соответствует значение $\bar{N}(E_0) = 0,98 \cdot 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. Это распределение получилось в результате дифференцирования sh-образной кривой зависимости Q_{ss} от E_F [уравнение (12.18)] в пренебрежении зависящими от температуры членами в правой части. Как видно из фиг. 37, даже вырожденное распределение поверхностных состояний может привести к заметной температурной зависимости заряда Q_{ss} ; sh-образное распределение нельзя распространить за интервал $\pm 0,13$ эВ около E_0 , иначе оно перекроет распределение собственных поверхностных состояний. Положение центра распределения поверхностных состояний E_0 — вопрос дискуссионный. Однако он не существен, коль скоро мы интересуемся *формой* распределения.

(Попытка непосредственного измерения V_D по методу фото-э. д. с. была безуспешной.) Согласно общепринятому мнению, для всех способов приготовления поверхности Si $0 < E_0 < -0,4$ эВ. Другими словами, поверхность относится к невырожденному p -типу. Франкль¹⁾ подверг сомнению указанное Алленом и Гобели максимальное число поверхностных состояний. Он указал, что, если не включать в рассмотрение вырожденные объемные примеси, это число станет меньше.

Остановимся теперь еще на некоторых работах, посвященных изучению свойств сколотого Si . Ван-Лаар и Шир [184—186] измеряли энергию Φ для различных концентраций примеси: от вырожденного n -типа до вырожденного p -типа. К сожалению, величина ϕ_w не измерялась, так что свойства поверхности пришлось определять из формы кривой Y от $h\nu$ для образцов p -типа. Из-за того что зоны изогнуты вверх, электроны в объеме (до $30 \text{ \AA}$ от поверхности) могут давать вклад в выход Y . Этот эффект изменяет форму кривых зависимости Y от $h\nu$. С помощью соответствующей математической обработки и подгонки кривой, а также привлекая данные этих же авторов по адсорбции индия, было найдено, что (вблизи E_F) $N \geq 10^{15} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. С помощью закона « $3/2$ », который соблюдается для выхода фотоэмиссии, эти авторы сделали заключение, что распределение поверхностных состояний должно быть sh -образным относительно нейтрального уровня E_0 . Этот уровень, как они утверждают, расположен вблизи центра запрещенной зоны ($E_0 = 0$). Рассчитанная таким образом плотность собственных поверхностных состояний составила $8 \cdot 10^{14} \text{ см}^{-2}$. В более ранней работе [184] эти же авторы, исходя из стабилизации уровня Ферми в пределах $\sim 0,1$ эВ при различных концентрациях примеси p -типа, получили нижний предел плотности донорных состояний $N \geq 2 \cdot 10^{13}$ в «узком энергетическом интервале». Ван-Лаар и Шир [184] связывали низкоэнергетические «хвосты» (подчиняющиеся кубическому закону или закону « $5/2$ ») с напряженными участками сколотого (практически сломанного) кристалла. Они привели одновременно

¹⁾ См. работу [8], стр. 201.

и другие данные, показывающие, как усиливается низко-энергетическая часть кривой фотоэмиссии с поверхности Si (111) при обдуве песком или травлении. Эта же закономерность наблюдалась также другими авторами [254, 255]. В своих последующих работах [185, 186] Ван-Лаар и Шир объясняли эту часть спектра эмиссией с поверхностных состояний, причем даже для образцов p -типа. Это определенно противоречит утверждению Аллена и Гобели [175—177]. Большинство других доказательств, рассматриваемых ниже, видимо, свидетельствуют в пользу предположения Аллена и Гобели, что поведение Φ (кубический закон) соответствует эмиссии из валентной зоны образцов p -типа.

Аспнес и Хандлер [178] исследовали изменение проводимости, происходящее после скола образцов Si p -типа, в одном случае легированных золотом, а в другом — бором. Определив степень увеличения или уменьшения проводимости после скола и сделав некоторые предположения о рассеянии, они нашли, что $E_0 = -0,27 \pm 0,08$ эВ, вне зависимости от предположений относительно Φ . Результат этих электрических измерений находится в хорошем согласии с данными Аллена и Гобели, но не с данными Ван-Лаара и Шири. Лучшим экспериментом был бы такой, в котором скалывались бы кристаллы с различной степенью легирования и был бы найден тот из них, проводимость которого *вовсе не изменялась* бы после скола. Это было сделано Хандлером [187]. Полученный результат для E_0 согласуется со значением Аллена и Гобели ($E_0 = -0,23$ эВ). Аспнес и Хандлер скалывали образцы при довольно низких температурах ($T = 105$ К), пытаясь выделить проводимость за счет поверхностных состояний. Перед сколом типичное значение проводимости равнялось $11 \cdot 10^{-12}$ См. Однако после скола проводимость не возрастала. Используя модели поверхностных состояний с непрерывным распределением уровней — слившейся или ch -образной зоны, — они вычислили подвижность и плотность состояний: $\mu_{ss} < 10^{-8}$ см²/(В·с), $\bar{N} > 1,8 \cdot 10^{12}$ см⁻²·эВ⁻¹. Ввиду полупроводникового характера поверхностных состояний этим авторам удалось объяснить экспериментальные результаты, исходя из *двухзонной* модели или модели

двух дискретных уровней. Ожидалось, однако, что при 300 К должна существовать заметная проводимость за счет поверхностных состояний, обусловленная термически активированными свободными носителями с плотностью $\sim 10^{12} \text{ см}^{-2}$ в зоне поверхностных состояний. Обнаружить ее не удалось. Таким образом, эти данные по проводимости для двух образцов оказались согласованными без введения дополнительной проводимости за счет поверхностных состояний. Оказалось, что с помощью одной только проводимости невозможно сделать выбор между распределениями поверхностных состояний. Дело представляется так, что поверхностные состояния в Si (а также в Ge, как это будет показано в п. 2) просто не проводят. Это находится в противоречии с основным предположением трехмерной теории поверхностных состояний. Причина такого противоречия не ясна. Может оказаться, что поверхностные состояния все-таки проводят, но соответствующий омический контакт осуществить не удалось.

Эксперименты с эффектом поля проводились Хенслером [188] и Хенслером и Хейландом [252] на кристаллах p -типа. Для защиты кристалла были приняты специальные меры, и эффект поля измерялся при различных расстояниях до электрода. Получены следующие результаты: $E_F = -0,36 \pm 0,02 \text{ эВ}$, $\mu_{fe} = -1 \text{ см}^2/(\text{В} \cdot \text{с})$ (p -тип) и $\Delta\sigma = (0,59 \pm 0,12) \cdot 10^{-6} \text{ См/ед. площади}$. Принимая модель *слившейся зоны* и учитывая (13.4), они получили $N = 1 \cdot 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. Но, как отмечалось в п. 2 § 13, те же результаты могут быть интерпретированы с помощью модели двух уровней с использованием соотношения (13.5). При этом получаются значения $2N = 1 \cdot 10^{14} \text{ см}^{-2}$ и $E_{sg} = 0,25 \text{ эВ}$. Последняя модель дает прекрасное согласие с предложенным Алленом и Гобели *распределением* собственных поверхностных состояний, равно как и с общепринятым положением нейтрального уровня поверхностных состояний E_0 ¹⁾. Аналогичные измерения с

¹⁾ В работе [333] для интерпретации оптических и электрических измерений поверхностных состояний на поверхности (111) Si была с успехом применена модель плотности состояний, аналогичная тем, которые были использованы Алленом и Гобели. — *Прим. перев.*

эффектом поля были выполнены Кавайи и Такишима [256] для Si *n*-типа. Они получили, что $E_F = -0,30$ эВ, но никаких оценок распределения поверхностных состояний сделано не было.

Хейнеман [216] измерял интенсивность сигнала ЭПР со сколотой поверхности монокристалла размером $0,7 \text{ см}^2$. Сигнал выделялся из фона с помощью метода аккумулярования. Этим способом была найдена полная спиновая плотность $0,8 \cdot 10^{14} \text{ см}^{-2}$. Это соответствует числу *неспаренных* электронов в поверхностных состояниях или в объемно-заряженной области, т. е. дает нижний предел. Хейнеман не обнаружил также сверхтонкой структуры и анизотропии у *g*-тензора. Он принял, что этот результат означает существование в поверхностных состояниях доли электронов, которые не локализованы в определенных местах и являются существенно свободными электронами. Были сделаны также следующие предположения (некоторые из них основываются на предшествующих работах [146, 238, 240]): 1) поверхностные атомы можно разделить на две группы; 2) некоторые из них поднимаются над поверхностью (их называют атомами *s*-типа); 3) другие опускаются (атомы *p*-типа); 4) эти две группы атомов расположены чередующимися рядами; 5) проводимость обусловлена только рядами с опущенными атомами и возможна лишь в продольном направлении; 6) ступеньки скола параллельны этим рядам и перпендикулярны контактам. Таким способом Хейнеман объяснил отсутствие проводимости [178] для зон поверхностных состояний. Он предположил, что распределение поверхностных состояний Аллена и Гобели (т. е. распределение на фиг. 34, а) может быть объяснено разделенными зонами, отвечающими поверхностным атомам *s*- и *p*-типа. Эта интерпретация неоднозначна. Другими возможными предположениями о происхождении этих распределений могут быть: 1) влияние эффекта Яна — Теллера, как было предложено Дакросом [257], или 2) спиновые корреляции поверхностных состояний, которые были рассмотрены Дэвисоном и Амосом [64], а также Томашеком [258]. В последнем случае зоны поверхностных состояний представляют собой нижнюю и верхнюю зоны, которым от-

вечают противоположно направленные спины ($\uparrow\downarrow$ соответственно). Чанг [219] изучал электронный парамагнитный резонанс на измельченном Si с различной степенью легирования. Оказалось, что наибольший сигнал наблюдается для образцов с собственной проводимостью.

Фишер [253] провел анализ энергетического спектра фотоэмиссии для кристаллов n -типа и вырожденного p -типа. При рассмотрении величина работы выхода ϕ_w не измерялась. Вместо этого в первую очередь изучались точки порога и насыщения энергетического спектра фотоэмиссии. Измеренная в интервале $80\text{ К} < T < 300\text{ К}$ температурная зависимость заряда оказалась слабой. Дальнейший анализ показал, что заряд Q_{ss} существенно не зависит от температуры в обоих образцах. Это свидетельствует об электронном вырождении и, как было показано, дает *нижний предел* величины $\bar{N}$ для поверхностных состояний в *двухзонной* модели: $\bar{N} \geq 4 \cdot 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. Это значение примерно на порядок меньше, чем полученное Алленом и Гобели $\bar{N} \geq 26 \cdot 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. Возвращаясь к фиг. 36 (см. также фиг. 35 и 37), мы видим, что отсутствие зависимости от температуры фактически имеет место только в *центрах* зон поверхностных состояний. Таким образом, это является более жестким ограничением, чем результат, получающийся из простых уравнений, использованных Фишером [уравнение (10) в его работе], а именно независимость от температуры гарантирует только то, что E_F лежит внутри зоны поверхностных состояний. *ch-образное* распределение, предложенное Алленом и Гобели, имеет, по-видимому, наименьшую температурную зависимость и, следовательно, лучше всего согласуется с результатами Фишера. Этот вывод получается из рассмотрения соотношения (12.17), а также из близости кривых, отвечающих $T = 300\text{ К}$ и $T = 75\text{ К}$ на фиг. 37. В хорошем согласии с работой Аллена и Гобели находятся такие результаты Фишера, как найденные положения уровней Ферми относительно краев зон у поверхности для образцов n - и p -типов.

Измерения с помощью эффекта поля, проведенные Пальмером и др. [259], показали, что $E_0 > 0$. Эта работа была впоследствии пересмотрена Хандлером [260].

который показал, что: 1) часть результатов вполне согласуется с результатами Аллена и Гобели и 2) другая часть ошибочна из-за большой емкостной связи с несклолотой поверхностью. Этим завершаются в основном работы, выполненные на склолоте Si.

Эксперименты со склолотыми и нагретыми поверхностями Si выполнялись Алленом и Гобели и были сообщены в их более поздней работе [203]. Однако измерения ϕ_w и Φ (в том числе энергетического распределения фотоэлектронов) проводились только при одном значении концентрации. Поэтому распределения поверхностных состояний найти не удалось. По-видимому, поверхностные состояния после нагрева участвуют в фотоэмиссии, так как потенциал Φ (кубический закон) относится к уровню Ферми, находящемуся внутри зоны поверхностных состояний. Было принято предположение, что $E_F = -0,23$ эВ, что совпадает со значением E_0 до нагревания. Предполагалось, что Φ в процессе нагрева должно уменьшиться на 0,23 эВ. Чанг и Хейнеман [217] изучали электронный парамагнитный резонанс в сверхвысоком вакууме на растолченном Si (предполагалось, что преимущественная огранка отвечала плоскостям {111}). Отжиг в сверхвысоком вакууме при температуре ~ 1100 К уменьшил сигнал ЭПР до $\sim 2\%$ по сравнению с сигналом от неотожженного растолченного Si. В некоторых случаях интерпретация этого результата предполагала возникновение новых поверхностных структур. Похожие результаты сообщались позднее Хейнеманом и др. [218] для растолченных на воздухе и механически полированных кристаллов Si. Интересно, что отжиг в сверхвысоком вакууме вызывает одновременно значительное *увеличение* фотоэмиссии с поверхностных состояний и сильное *уменьшение* сигнала ЭПР.

Кремний, подвергнутый бомбардировке и отжигу, был исследован Алленом и Гобели [203], Лоу [261], Хейландом и Ламатшем [262], Нестеренко и Снитко [192, 193], а также Диллоном и Фарнворсом [263]. Несмотря на столь значительные усилия, распределение поверхностных состояний осталось невыясненным. Произошло это в первую очередь потому, что концентрация примесей оставалась неизменной. Аллен и Гобели, подобно своему

результату для сколотого и отожженного образца, получили $E_0 \sim -0,23$ эВ. Однако бомбардированная и отожженная поверхность имеет следующие отличия: 1) отсутствует линейный участок графика зависимости Y от $h\nu$ (вместо него наблюдался только искривленный участок); 2) образуется тонкий поверхностный слой p -типа (по-видимому, из-за поверхностного загрязнения). Он образуется, даже если ионная бомбардировка предшествует первому нагреву. Это противоречит утверждению Диллона и Омана [264]¹⁾, которые нашли, что пленка p -типа может быть удалена бомбардировкой перед первым нагревом. В сколотых и отожженных образцах, например, такая пленка не образуется. Как будет видно из дальнейшего, пленка p -типа представляет собой источник затруднений для многих исследователей, применяющих метод бомбардировки и отжига.

Диллон и Фарнsworth [263] измерили работу выхода ϕ_w и потенциал Φ для поверхностей (111), (110) и (100) Si p -типа. Была найдена следующая последовательность значений работы выхода:

$$\phi_w(100) > \phi_w(110) > \phi_w(111).$$

В таком же порядке расположены поверхностные плотности разорванных связей, отнесенные к соответствующим граням (табл. 2, § 9) [136]. Совершенно неожиданно оказалось также, что порог фотоэмиссии Φ меньше ϕ_w . Авторы связывают это с наличием примеси кислорода в образцах. Таким образом, вывести заключение о распределении поверхностных состояний из этих данных невозможно. Лоу [261], будучи осведомлен о проблеме загрязнения и образования пленки p -типа, принял соответствующие предохранительные меры. Проводя одновременные измерения проводимости (характеристик контакта) $p-n$ -перехода и фотопроводимости, он обнаружил, что $E_F = 0,37$ эВ как для n -, так и для p -области. В присутствии пленки p -типа E_F должно было бы находиться ближе к центру щели ($E_F = 0$).

Хейланд и Ламатш [262] также знали о проблеме загрязнения и приняли соответствующие меры. Они

¹⁾ См. также [7], стр. 441.

исследовали свойства проводимости Si вдоль нити *p*-типа с сопротивлением 1000 Ом·см (круговое сечение содержало различные кристаллические плоскости). С помощью чередования бомбардировки и отжига им удалось получить минимум поверхностной проводимости. Ясно, что бомбардировка без отжига сохраняет поверхность *n*-типа. Было найдено, что $E_F = -0,25$ эВ. Измерения подвижности в эффекте поля дали $\mu_{fe} = -0,01$ см²/(В·с) и $\bar{N} = (1,5 \pm 0,5) \cdot 10^{15}$ см⁻²·эВ⁻¹ в точке E_F , если принимается модель слившейся зоны, или $N \geq 10^{14}$ см⁻² для модели одиночного дискретного уровня, расположенного в точке E_F . Таким образом, эти результаты подтверждают вывод Аллена и Гобели, что энергия E_F для скелотого и нагретого или бомбардированного и отожженного Si находится внутри зоны поверхностных состояний, причем порог фотоэмиссии обусловлен фотоэмиссией с поверхностных состояний. Нестеренко и Снитко [192], а также Нестеренко с соавторами [193] измеряли импульсный эффект поля и подвижность в эффекте поля для поверхности (111) образцов *n*- и *p*-типа. В качестве прослойки использовалась пластинка слюды, контакт которой с поверхностью Si предотвращался присутствием невытравленных выступов на поверхности Si. Результаты были согласованы с двухзонной моделью, представленной на фиг. 36. Из динамики процессов ясно, что верхняя зона (начинающаяся в точке $E_1 \sim 0,1$ эВ)—зона электронных акцепторных состояний, а нижняя ($E_2 \sim 0,1$ эВ)—зона дырочных акцепторных (электронных донорных) состояний. Предполагается, что нейтральный уровень находится в центре запрещенной зоны ($E_0 = 0$). Далее, $N_1 = N_2 = 10^{14} - 10^{15}$ см⁻² и $\bar{N}_1 = \bar{N}_2 = 1 \cdot 10^{15}$ см⁻²·эВ⁻¹. Заметим здесь, однако, следующее. Никакого эффекта, вызванного адсорбированным монослоем O₂, Au или Sb, найдено не было. Это может быть связано с присутствием изолирующего окисного слоя на поверхности Si, не удаленного полностью с помощью примененной авторами процедуры (она была, вероятно, той же, что и у Фарнворса).

Как отмечалось ранее, большое число экспериментов, выполненных на чистом Si, включая и точечную

холодную эмиссию [265], дали относительно бедную информацию о распределении поверхностных состояний. Основная причина этого состоит в том, что большинство экспериментов было выполнено для сравнительно узкого интервала концентраций примеси. В кремнии приложенное поле просто не может достаточно сильно «раскачать» зоны ввиду закрепляющего действия поверхностных состояний. Также встречается затруднение получения воспроизводимых поверхностей Si. Причины этого следующие: 1) трудность удовлетворительного скола Si; 2) склонность к образованию примесного слоя p -типа у бомбардированного и отожженного Si; 3) образование различных поверхностных фаз перестройки, зависящих от обработки.

Теоретические исследования, относящиеся к трехмерной задаче о поверхностных состояниях в алмазе, Si и Ge, были выполнены Коутецким и Томашеком [26], Чейвзом и др. [132], Джонсом [134], Коутецким [266], Пью [267] и Бартошем [268]. Ниже мы кратко коснемся этих работ и их результатов.

Коутецкий и Томашек [26] выбрали для описания объемных свойств кристалла типа алмаза метод сильной связи. В качестве атомных орбиталей выбирались sp^3 тетраэдрические орбитали. Образование поверхности (111) достигалось мысленным удалением атомов (т. е. требованием исчезновения соответствующих коэффициентов разложения) при неизменных поверхностных кулоновских и резонансных интегралах. Таким образом, разыскивались так называемые шоклиевские состояния. Были найдены: 1) объемная зонная структура, содержащая зону проводимости, валентную зону и две валентные зоны, энергии которых не зависят от k ; 2) *слившаяся зона* поверхностных состояний, как на фиг. 34, в, ширина которой E_b значительно меньше, чем объемная запрещенная зона $E_b \leq 0,2 E_g$; 3) положение нейтрального уровня E_0 на некотором расстоянии выше центра щели и его зависимость от выбора параметров.

Впоследствии Коутецкий [266] рассмотрел эффекты возмущения поверхностного кулоновского интеграла (т. е. таммовские состояния). Если это возмущение достаточно велико, слившаяся зона поверхностных состояний

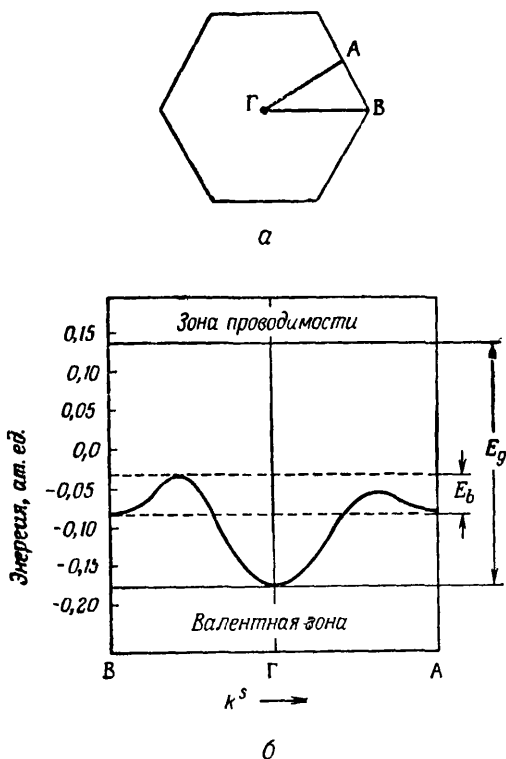
может сместиться ниже центра запрещенной зоны. Позднее было рассмотрено и некоторое возмущение поверхностного резонансного интеграла. Это также сдвигает зону поверхностных состояний в нижнюю часть запрещенной зоны (E_0 становится отрицательным). В этих работах указывалась только ширина зоны поверхностных состояний, а не закон дисперсии $E(\mathbf{k})$ или плотность состояний.

Пью [267] провел более тщательный расчет, используя те же sp^3 -орбитали в качестве исходных состояний. Были рассчитаны некоторые из 33 необходимых двухатомных молекулярных интегралов между первыми, вторыми и третьими ближайшими соседями в поверхностной области и в объеме. Он также рассчитал закон дисперсии $E(\mathbf{k})$ для двумерной гексагональной поверхностной зоны (фиг. 38, а и б). Минимум при этом оказался в точке Γ , а максимумы не попали в симметричные точки. Ширина объемной запрещенной зоны равна E_g . Кроме того, Пью рассчитал *слившуюся зону* поверхностных состояний для поверхности (111) алмаза, лежащую ниже центра запрещенной зоны. Причем около 90% всех состояний было сосредоточено в очень узком интервале, обозначенном через E_b на фиг. 38, б.

Чейвз и др. [132] с помощью $\mathbf{k} \cdot \mathbf{p}$ -метода получили точную картину объемной зонной структуры Si и Ge. Они рассчитали энергии поверхностных состояний в отдельных точках зоны Бриллюэна, принимая ступенчатый разрыв потенциала в середине расстояния между атомными слоями. Эти результаты показывают, что поверхностное состояние в точке X_1 зоны Бриллюэна должно находиться чуть ниже (на 0,07 эВ) края зоны проводимости в Si и немного выше — в Ge.

Бартош [268] расширил применение $\mathbf{k} \cdot \mathbf{p}$ -метода, исследовав поверхности (100) Si и Ge с различным положением точки разрыва поверхностного потенциала. Он получил, что в точке Γ двумерной зоны Бриллюэна поверхностное состояние находится примерно на 0,20 эВ выше центра запрещенной зоны в случае Ge и совершенно отсутствует у Si. Эти результаты оказались явно нечувствительными к положению точки разрыва.

Джонс [134]¹⁾ выполнил расчет для поверхности (110) Si с помощью метода псевдопотенциала и построил диаграмму $E(k)$ зоны поверхностных состояний

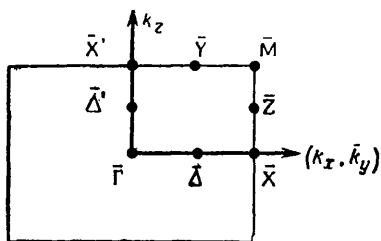


Фиг. 38. Теоретическая зонная структура поверхностных состояний на неперестроенной плоскости (111) алмаза, вычисленная Д. Пью [267].

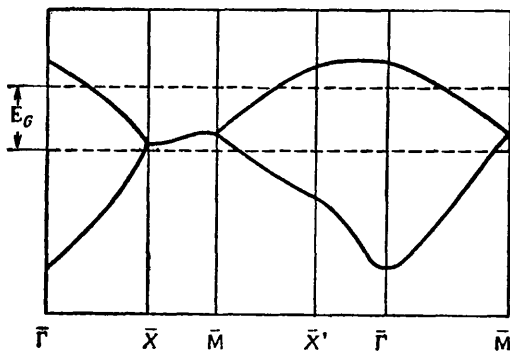
(фиг. 39). Он нашел сплошную зону поверхностных состояний, перекрывающую всю энергетическую щель, подобно схеме на фиг. 34, e, причем многие из состояний находятся вне запрещенной зоны.

¹⁾ См. также работу [336], где были рассмотрены поверхности (100) и (111). — Прим. перев.

В рассмотренных выше теориях не принималась во внимание поверхностная перестройка, на существование которой, как известно, указывают эксперименты по



a



б

Фиг. 39. Теоретическая зонная структура поверхностных состояний на неперестроенной плоскости (110) Si, вычисленная Р. О. Джонсом [269].

ДЭНЭ на поверхностях Si и Ge. Не удивительно поэтому, что сравнение теории поверхностных состояний с экспериментом не обнаруживает в настоящее время заметного соответствия. Так, например, некоторые рассчитанные теоретически поверхностные состояния находятся вблизи зоны проводимости, что не согласуется с экспериментом. Далее эти расчеты не дают *парных* донорно-акцепторных состояний с отрицательным значением E_0 , как показано на фиг. 34, а, б и г,

2. Германий

Распределение поверхностных состояний для Ge все еще остается спорным, несмотря на то, что этой проблемой занималось большое число исследователей. Это объясняется, вероятно, тремя основными причинами. Во-первых, поверхностные состояния часто оказываются расположенными вблизи края валентной зоны и во многих случаях попадают *внутрь* валентной зоны. Это создает трудности для их выявления и анализа. Во-вторых, Ge имеет непрямую запрещенную зону шириной 0,67 эВ, что составляет около половины запрещенной зоны Si. В силу этого обстоятельства закрепление уровня Ферми в Ge может быть обеспечено с помощью гораздо меньшей плотности поверхностных состояний [270]. Иначе говоря, для одинаковых заданных плотностей поверхностных состояний в Ge и Si закрепление уровня Ферми в Ge будет значительно сильнее, чем в Si. Похоже, что в основном дело именно в этом. В-третьих, подавляющее большинство работ относится к измерениям подвижности в эффекте поля. Как отмечалось выше в разделе, посвященном Si, этот подход к анализу распределения поверхностных состояний не вполне удовлетворителен, до тех пор пока одновременно не будут измерены другие параметры.

Картина ДЭНЭ для Ge очень похожа на картину для Si [52, 234, 235]. В частности, скол при комнатной температуре дает перестроенную поверхность Ge (111)- $1 \times \sqrt{3}$. После нагревания до 470 К она переходит в Ge (111)-8. Дальнейший нагрев выше 570 К дает структуру Ge (111)-12 [60]. Поверхность (111) после бомбардировки и отжига дает ту же картину ДЭНЭ, что и поверхность после скола и нагрева: Ge (111)-12. Имеются, однако, тонкие различия, которые обнаружил Фарнsworth, применяя в методе ДЭНЭ при получении дифракционных картин клетку Фарадея. С помощью этого метода Фарнsworth с сотрудниками [251] показал, что поверхность, подвергнутая бомбардировке и отжигу, в большей степени идеальна, чем сколотая и нагретая. Поверхности (100) и (110) дают свои собственные

картины ДЭНЭ-рефлексов (для них нет еще установившихся обозначений).

Рассмотрим сначала свойства сколотого Ge, исследовавшиеся Гобели и Алленом [270], Хенслером [60, 271] и Пальмером с сотрудниками [222, 272]. Гобели и Аллен пытались определить распределение поверхностных состояний в Ge, следуя той же процедуре, которую они с успехом применяли для Si. Они измерили ϕ_w и Φ (экстраполяция к кубическому закону) для различных концентраций примеси от вырожденного p -типа до вырожденного n -типа. Получился удивительный результат: $\phi_w = \Phi$ независимо от степени легирования в пределах точности эксперимента $\pm 0,02$ эВ. Таким образом, уровень Ферми очень сильно закреплен на поверхности. Он может быть смещен самое большее на $0,04/0,67 \approx 6\%$ ширины запрещенной зоны. С другой стороны, в Si сдвиг составляет $0,2/1,1 \approx 20\%$ ширины запрещенной зоны. Было рассмотрено несколько возможных моделей объяснения этих данных. При имеющейся информации лучшей была признана модель *двух дискретных уровней*, центр распределения которых находится на краю валентной зоны:

$$E_0 = -0,33 \pm \text{несколько } kT, \quad E_1 = E_0 + \text{несколько } kT,$$

$$E_2 = E_0 - \text{несколько } kT, \quad E_{sg} \sim \text{несколько } kT,$$

$$N_1 = N_2 \geq 1,4 \cdot 10^{13} \text{ см}^{-2}.$$

Таким образом, по крайней мере $1/50$ часть ожидаемого числа собственных поверхностных состояний в Ge ($7 \cdot 10^{14} \text{ см}^{-2}$) объясняется этим неравенством. Рассмотренное распределение напоминает распределение в Si, за исключением двух различий: 1) в Ge нейтральный уровень E_0 находится практически на краю валентной зоны, тогда как в Si он расположен примерно на 0,3 эВ выше края валентной зоны; 2) щель E_{sg} в Ge значительно уже, чем в Si. Заполнение поверхностных состояний Q_{ss} в Ge напоминает при этом кривую, отвечающую 75 К, в модели *двух дискретных уровней* (фиг. 35). (Характер заполнения уровней отвечает примерно соотношению $E_{sg}/kT = \text{const.}$) Из фиг. 35 можно видеть, что, когда зонная щель составляет всего только не-

сколько kT , выявить структуру поверхностных состояний по величине Q_{ss} довольно трудно. Действительно, кривая зависимости Q_{ss} от E_F сильно напоминает кривую при 300 К модели *слившейся зоны* (фиг. 37). Выделить структуру поверхностных состояний на фоне «теплового шума» можно только после изучения низкотемпературного поведения Q_{ss} .

Чтобы окончательно решить, связано ли поведение энергии Φ (кубический закон) с эмиссией из поверхностных состояний или из валентной зоны, были проведены дополнительные эксперименты с использованием поляризованного света [273, 274]. Была обнаружена индуцированная поляризация распределения эмиттированных электронов, сохраняющаяся и в области кубического «хвоста». Это означает, что рассеяние является *не* полностью диффузным и пороговые процессы должны быть рассмотрены как прямые в противоположность первоначальному утверждению, сделанному для Si [175—177]. Такие процессы были названы «квазипрямыми». В них при поглощении $k_{\parallel}$ все еще сохраняется в противоположность $k_{\perp}$. Остается не вполне ясным тот факт, что эксперимент показывает только приблизительно 40%-ный вклад прямых процессов (без рассеяния) в области «хвоста». Таким образом, другие процессы рассеяния, которые тоже дают вклад, не могут быть отброшены. Аллен и Гобели указали на возможность (хотя это и маловероятно) того, что поверхностные состояния обуславливают указанный выше поляризационный эффект. Таким образом, проблема происхождения порога фотоэмиссии Φ еще не решена окончательно, несмотря на использование поляризационных экспериментов.

Хенслер [60] измерил проводимость и подвижность в эффекте поля для шести образцов Ge с одинаковой примесью p -типа. Он получил $\mu_{fe} = -100 \text{ см}^2/(\text{В} \cdot \text{с})$. Это показывает, что поверхность относится к p -типу в соответствии с результатом Аллена и Гобели. Однако положение уровня Ферми у поверхности оказалось равным $E_F = -0,23 \text{ эВ}$ (невырожденный тип). Хенслер также вычислил $\bar{N}$ в точке E_F , исходя из модели *слившейся зоны*, и нашел, что $\bar{N} = 2 \cdot 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. Как

обычно, в исследованиях подобного рода необходимо ввести предположения о механизме поверхностного рассеяния. Хенслер использовал данные по подвижности, полученные Хандлером и Эйзенхауэром [181] из измерений магнитосопротивления на поверхностях Ge, приготовленных с применением бомбардировки и отжига.

Пальмер и др. [222] пытались измерить распределение поверхностных состояний, используя сколотые $p-n-p$ -переходы и исследуя канальный эффект. Было найдено, что управляющее напряжение не может повлиять на свежесколотые образцы ввиду того, что избыточная плотность дырок, обусловленная акцепторными поверхностными состояниями, оказывается много больше, чем изменение числа понижированных внешним полем доноров. Другими словами, плотность поверхностных состояний слишком велика, чтобы ее можно было измерить. Было, однако, сделано заключение, что $N > 1,5 \cdot 10^{12} \text{ см}^{-2}$. Варьируя дозу кислорода, этим авторам удалось уменьшить Δp , уменьшить число поверхностных состояний и измерить их распределение в интервале энергий, покрывающем запрещенную зону. Для этого использовались туннельные вольт-амперные характеристики. Эти наблюдения и сам метод интересны, но практически непригодны для исследования поверхностных состояний на чистом Ge. В продолжение своих предыдущих работ указанные авторы [272] изучали сколотые образцы Ge n - и p -типа. Результаты получились фактически одинаковыми. Была измерена подвижность в эффекте поля с использованием одной из плоскостей сколотого образца в качестве электрода. Она помещалась на расстоянии всего лишь около 0,4 см от поверхности. Полученное значение $\mu_{fc} = -(50 - 150) \text{ см}^2/(\text{В} \cdot \text{с})$ оставалось неизменным по крайней мере в течение 100 ч при давлении 10^{-9} мм рт.ст. Оценка величин E_F и N дала следующие результаты: $E_F = -0,10 \pm 0,02 \text{ эВ}$ и $N \geq 10^{11} \text{ см}^{-2}$. В обеих работах [222, 272] наблюдалось, однако, резкое возрастание давления (до $10^{-8} - 10^{-7}$ мм рт.ст.) после скола (и, вероятно, еще больше вблизи кристалла). Как признает автор, это усложняет интерпретацию результатов, относящихся к чистым поверхностям.

Банбьюри с соавторами [275], а также Барнес и Банбьюри [276] получили результаты, заметно отличающиеся от упомянутых выше. Так, они нашли, что $E_F = -0,05$ эВ. Критика этой работы была дана Мейни [7] и мотивировалась существованием значительной емкости между частями сколотого кристалла [276] и присутствием прослойки слюды, находящейся в тесном контакте с изучаемой поверхностью [275].

Важные эксперименты со сколотым и нагретым Ge были выполнены Хенслером [60]. Он рассчитал плотность состояний $\bar{N}$ из подвижности в эффekte поля как функцию температуры нагрева T вплоть до 650 К. Результаты были обобщены в виде кривой зависимости $\bar{N}$ от T , которая показала монотонное уменьшение $\bar{N}$ от значения $2 \cdot 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$ при $T = 300$ К до $2 \cdot 10^{13} \text{ см}^{-2} \cdot \text{эВ}^{-1}$ при $T = 650$ К. Величина $\bar{N}$ относится, конечно, к уровню Ферми у поверхности E_F , который при нагревании повышается от $-0,23$ до $-0,15$ эВ. Наиболее удивительным результатом является отсутствие собственной структуры (т. е. локальных максимумов или минимумов) у зависимости $\bar{N}$ от T . Результат объяснялся тем, что измеренные поверхностные состояния не сосредоточены у границ зерен или каких-либо других дефектов. Обоснование этого утверждения заключается в следующем: 1) Из экспериментов по ДЭНЭ при различных температурах найдено, что в некоторых интервалах температуры поверхность образует стабильные фазы, однако в момент структурного перехода она сильно разупорядочена. 2) Если измеренные поверхностные состояния обусловлены разупорядочением у границ зерен или другими дефектами, то характеристики этих состояний должны иметь максимум при структурных превращениях. 3) На самом деле это не наблюдается. Хенслер, рассматривая возможные причины появления поверхностных состояний у Ge, делает следующее заключение на основании своих экспериментов: «обрыв кристаллического потенциала у поверхности» должен быть отвергнут в пользу «особого расположения атомов в приповерхностных слоях, отличного от их расположения в объеме».

Недавно Хенслером [271] были выполнены эксперименты по измерению подвижности в эффекте поля для сколотого Ge после различных стадий отжига в интервале температур от 300 до 600 К. Одновременно регистрировалась картина ДЭНЭ в вакууме при давлении менее $5 \cdot 10^{-10}$ мм рт.ст. Ранее указанное им соотношение [60] между поверхностной проводимостью и подвижностью в эффекте поля для различных температур уверенно подтвердилось. Детальный анализ данных был проведен на основе различных моделей, включая модель *двух дискретных уровней* (фиг. 34, а) и модель *сливающейся зоны* (фиг. 34, в). Введение по крайней мере двух распределений обусловлено тем, что данные по ДЭНЭ показывают постепенный переход от исходной структуры 2×1 к фазе 2×8 при более высоких температурах. Путем подгонки кривых для выбранных простых моделей Хенслер нашел, что рассмотренные им распределения (фиг. 40) наилучшие. Сразу после скола образуется система двух дискретных уровней, отвечающих структуре 2×1 . При этом

$$E_0 = -0,30 \text{ эВ}, \quad E_1 = -0,36 \text{ эВ}, \quad E_2 = -0,21 \text{ эВ}, \\ E_{sg} = 0,15 \text{ эВ}, \quad 2N = 0,6 \cdot 10^{14} \text{ см}^{-2}.$$

После отжига при 150°C появляется другая система двух дискретных уровней. Она соответствует структуре 2×8 (точнее будет назвать ее «слабой 8-структурой»). Здесь

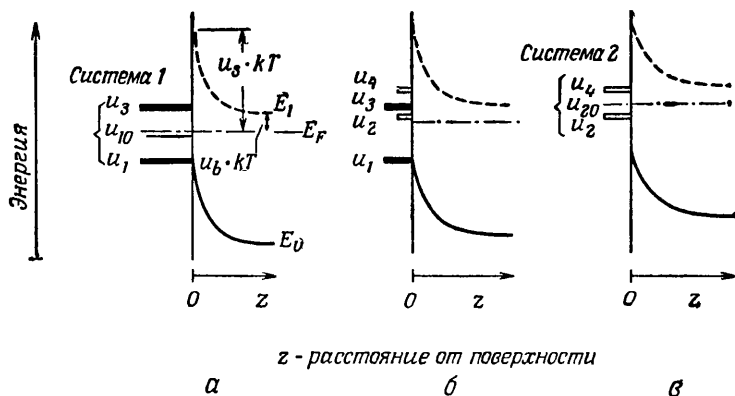
$$E_0 = -0,21 \text{ эВ}, \quad E_1 = -0,25 \text{ эВ}, \quad E_2 = -0,17 \text{ эВ}, \\ E_{sg} = 0,08 \text{ эВ}, \quad 2N = 0,3 \cdot 10^{13} \text{ см}^{-2}.$$

При промежуточных температурах отжига имеются обе системы. Полученные значения N на самом деле лишь нижние пределы, поскольку состояния, находящиеся далеко от E_1 и E_2 , не дают вклада из-за принципа Паули. Эти результаты, наконец, согласуются с недавно полученными оптическими данными по поверхностным состояниям у сколотых поверхностей Ge [210].

Гобели и Аллен [270] измерили ϕ_w и Φ для одного сколотого и нагретого кристалла p -типа. Оказалось, что величина E_F изменяется от $-0,33$ эВ перед нагревом до

значения $-0,22$ эВ после нагрева. Общая тенденция к образованию более собственной (хотя все еще p -типа) поверхности после нагрева согласуется с данными Хенслера [60, 271]. На этом завершаются эксперименты со сколотым, а также сколотым и нагретым Ge.

Эксперименты с Ge, подготовленным бомбардировкой и отжигом, выполнялись Алленом и Фаулером [137],



Фиг. 40. Распределение поверхностных состояний для плоскости (111) сколотого Ge по Хенслеру [271].

а — структура 2×1 после скола; *б* — промежуточная структура после нагрева до 80°C ; *в* — слабая структура 2×8 после нагрева до 150°C .

Форманом [277], Хандлером и Портным [278], Марго-нинским [194], Нестеренко и Снитко [191], а также Кобаяши с соавторами [179, 220]. Их результаты до некоторой степени противоречивы, за исключением общего вывода, что все поверхности относятся к p -типу (возможно, вырожденному) и что поверхность (100) имеет более четко выраженный p -тип, чем поверхность (111). Аллен и Фаулер [137] еще в 1957 г. измерили ϕ_w при четырех значениях концентрации n - и p -примесей для поверхностей (111), (110) и (100). Отмечалось, что очистка поверхности бомбардировкой ионами Ag является быстрой и эффективной. Однако она не обеспечивает сохранения чистоты поверхности при последующем нагреве, если тщательно не удалены газы из объема образца. Была также рассмотрена возможность

образования в процессе отжига пленок загрязнения на поверхности. Источником этого могут быть следы газовых примесей в Ag или на стенках камеры. Доказательством того, что в данных экспериментах загрязнение *не* имело места, было постоянство величины ϕ_w при медленном охлаждении образца от температуры отжига 1000 К. Пленки загрязнения *p*-типа также отсутствовали, поскольку образцы *n*-типа оставались такими же. Часть разногласий между авторами (см. ниже) могла быть обусловлена недостатками контроля этих усложняющих факторов. Результаты Аллена и Фаулера [137] показали, что $\phi_w(100) > \phi_w(110) > \phi_w(111)$ и $\phi_w(100) - \phi_w(111) = 0,06$ эВ. Эти значения нечувствительны к легированию, что согласуется с более поздними результатами [270]. Нечувствительность к температуре в процессе охлаждения, а также эксперименты с фото-э.д.с. Гаррета и Браттейна [279] позволили сделать заключение, что изгиб зон у поверхности $eV_D > 0,1$ эВ и что для плотности поверхностных состояний должно быть $\bar{N} > 10^{14}$ см⁻²·эВ⁻¹. Обращаясь, наконец, к теоретическим исследованиям Смолуховского [280], основанным на представлении о сглаживании волновых функций электронных состояний у поверхности, авторы отметили, что получающийся порядок значений работы выхода $\phi_w(111) > \phi_w(110) > \phi_w(100)$ противоположен наблюдаемому экспериментально. Был сделан вывод, что ненасыщенные связи и возмущение поверхности, не учтенные Смолуховским, ответственны за это противоречие.

Хандлер [136] в прекрасной обзорной статье (по работам до 1958 г.), посвященной Ge, подготовленному бомбардировкой и отжигом, развил анализ Аллена и Фаулера [137] несколько дальше. Он заметил, что последовательность $\phi_w(100) > \phi_w(110) > \phi_w(111)$ та же, что и последовательность поверхностной плотности ненасыщенных связей (табл. 2, § 9). Вопрос о такой интерпретации в настоящее время остается открытым ввиду перестройки, выявленной в последующих экспериментах по ДЭНЭ. Исходя из результатов многочисленных экспериментов, выполненных им самим и другими авторами в 1955—1956 гг., Хандлер получил оценки $N > 1 \cdot 10^{13}$ см⁻² и $\bar{N} \sim 4 \cdot 10^{14}$ см⁻²·эВ⁻¹.

Форман [277] измерил подвижность в эффекте поля и проводимость на поверхностях (111) и (100) образцов n -типа. Он нашел, что поверхность (100) имеет более резко выраженный p -тип, чем поверхность (111). Это не противоречит описанным выше измерениям ϕ_w . Таким образом, если, как это вычислил Форман, поверхность (111) характеризуется уровнем Ферми, находящимся на расстоянии не более kT от края валентной зоны, то поверхность (100) должна относиться к еще более вырожденному p -типу. Эти утверждения были, однако, подвергнуты критике со стороны Франкля [190] на том основании, что непротиворечивость результатов требует либо меньшей, чем единица, вероятности зеркального отражения от поверхности, либо присутствия наряду с акцепторными некоторого числа донорных состояний.

Франкль [190] сделал вывод, что подходящим распределением поверхностных состояний для поверхности (111) служит модель *слившейся зоны*, подобная показанной на фиг. 37, но бесконечно узкой. Оценка параметров этого распределения дала: $E_0 = -0,25$ эВ при $2N = 2 \cdot 10^3$ см⁻². Это существенно меньше, чем ожидаемая плотность поверхностных состояний $2N \sim \sim 7 \cdot 10^{14}$ см⁻². Следует подчеркнуть, что аргументы Франкля основывались на результатах по адсорбции кислорода на поверхности Ge. Поэтому вопрос об их применимости к *чистому* Ge остается открытым.

Измерения, подобные проведенным Форманом, были независимо выполнены Хандлером и Портным [278]. Кроме уже известных, были получены также результаты в температурном интервале $77 \text{ К} < T < 300 \text{ К}$ для поверхности (100) образца n -типа. Они нашли, что свойства не зависят от температуры, и сделали заключение, согласующееся с данными Формана, что поверхность (100) должна относиться к вырожденному p -типу. Результаты интерпретировались с помощью *двухзонной* модели поверхностных состояний, наподобие изображенной на фиг. 34, б. При этом энергия E_1 находилась глубоко внутри валентной зоны. Отсутствие температурной зависимости μ_{fe} и $\Delta\sigma$ позволило получить *верхний предел* плотности поверхностных состояний

$N \leq 4 \cdot 10^{13} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. [Предположение авторов о квантовании объемных дырок (названных впоследствии «легкими дырками») в сильно заряженном слое не связано с указанными выводами.] Франкль [189] применил различные подходы при интерпретации этих данных. Вычисляя девять различных поверхностных свойств при температурах 300, 200 и 100 К, ему удалось показать (в предположении полностью зеркального отражения), что эти данные согласуются со значением $E_0 = -0,25 \text{ эВ}$. Поверхность относится к p -типу, но не вырождена. Принимая модель *слившейся зоны* нулевой ширины, Франкль получил $N \geq 1,2 \cdot 10^{13} \text{ см}^{-2}$.

Маргонинский [194] также выполнил аналогичные эксперименты. Он измерил подвижность в эффекте поля, поверхностную проводимость и скорость поверхностной рекомбинации для поверхностей (111) и (100) образцов n -типа. Им было рассчитано (при условии полностью зеркального отражения), что поверхность (111) явно относится к p -типу, но не вырождена, а поверхность (100) — к вырожденному p -типу. Точнее говоря, для поверхности (111) $E_F = -0,25 \text{ эВ}$ и $\bar{N} = 2,4 \cdot 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$, а для (100) $E_F = -0,35 \text{ эВ}$ и $\bar{N} = 5 \cdot 10^{14} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. Эксперименты по поверхностной рекомбинации показали, что вблизи E_F имеются поверхностные центры рекомбинации акцепторного типа. Их общая плотность $N \sim 10^{14} \text{ см}^{-2}$. Утверждается, что поверхностные центры могут быть связаны с поверхностными состояниями. Об этом свидетельствуют эксперименты по измерению проводимости. Все это в целом согласуется с более ранними измерениями поверхностной рекомбинации Уанга и Уоллиса [281].

Сравнительно недавно Нестеренко и Снитко [191] выполнили подобные эксперименты для поверхности (111) образца p -типа. Поверхность оказалась относящейся к сильно вырожденному p -типу. Для объяснения этого результата применялась та же модель распределения поверхностных состояний, которую приняли Хандлер и Портной [278]. В результате получились значения $E_F = -0,38 \text{ эВ}$ и $\bar{N} = 10^{15} \text{ см}^{-2} \cdot \text{эВ}^{-1}$. Отличие от результатов других авторов может быть вызвано тем, что поверхность Ge готовилась бомбардировкой ионами

гелия, а не аргона, а также тем, что для изоляции электрода от поверхности Ge использовались кварцевые нити. Полученное значение подвижности $\mu_{fe} = 1400 \text{ см}^2/(\text{В} \cdot \text{с})$ примерно на порядок больше, чем значения, измеренные другими авторами, и приближается к величине объемной подвижности $\mu_b = 1900 \text{ см}^2/(\text{В} \cdot \text{с})$. Были также измерены спектры шумов. (Они оказались наименьшими в случае свободной поверхности и увеличивались в десять раз после окисления.)

Кобаяши с соавторами [179, 220] пытались обнаружить проводимость за счет поверхностных состояний того типа, который был найден в случае Si Аспнесом и Хандлером [178]. Были измерены проводимость σ и эффект Холла для поверхности (111) Ge в интервале $1,7 \text{ К} < T < 300 \text{ К}$, однако взятые ими образцы были «термически инвертированы». Найденная проводимость не исчезала при самых низких температурах и была, естественно, связана с проводимостью двумерной зоны поверхностных состояний. Образцы приготавливались нагревом без бомбардировки и отжига. В следующем году Кобаяши с соавторами [179] провели те же эксперименты, но на этот раз после бомбардировки и отжига. Оказалось, что с понижением температуры проводимость исчезает. Предыдущие «аномальные результаты» [220] были приписаны примесям p -типа, «вероятно бору, перешедшему со стеклянных деталей».

Как видно из предыдущего обсуждения, большинство наблюдаемых свойств поверхностных состояний в Ge в настоящее время противоречивы. Табл. 3 приведена здесь, чтобы подчеркнуть это утверждение. В первой колонке дана классификация поверхностей Ge по способу их приготовления: сколотая поверхность (111), сколотая и нагретая (111), бомбардированная и отожженная (111) и бомбардированная и отожженная (100). Во второй колонке перечислены ссылки на экспериментальные работы и соответствующие комментарии. Три оставшиеся колонки отвечают вырожденному p -типу, сильно вырожденному и невырожденному типам поверхности. Каждой ссылке соответствует крестик в одной из этих колонок. Рассмотрение таблицы показывает заметный

Таблица 3

Типы чистых поверхностей Ge

Приготовление поверхности	Литературный источник	Вырожденный p -тип	Сильно вырожденный p -тип	Невырожденный p -тип
Скол (111)	[270]	×		
	[60, 282] *			×
	[259]			×
	[7]	×		
Скол и нагрев (111)	[60, 282] *			×
	[270]			×
Бомбардировка и отжиг (111)	[277]	×		
	[190, 8]			×
	[194]			×
	[191]		×	
	[7]	×		
Бомбардировка и отжиг (100)	[277]		×	
	[278]		×	
	[8, 189]			×
	[194]	×		
	[7]		×	

* См. также *Henzler M.*, *Phys. Rev.* (будет опубликовано).

разброс результатов. Количество крестиков в трех колонках равно соответственно 5, 4 и 7. По-видимому, статистической корреляции здесь нет (возможности равновероятны). Другой неприятный факт состоит в том, что, несмотря на подобие дифракционных картин (метод ДЭНЭ), полученных со сколотых и нагретых, а также с бомбардированных ионами и отожженных поверхностей (111) Ge [52, 234, 235], таблица наличия такого подобия не подтверждает.

Если поверхность Ge относится к вырожденному или сильно вырожденному p -типу (согласно таблице вероятность этого $\sim 50\%$) или если E_1 находится близко от края валентной зоны, то это может иметь важные теоретические следствия. 1) Должны существовать поверх-

ностные состояния, попадающие внутрь валентной зоны. 2) Эти состояния должны быть *локализованными* (точнее резонансными). 3) Теоретические соображения в трехмерной задаче § 8 относительно отсутствия пересечения зон в k -пространстве могут объяснить возникновение этих состояний. 4) Существует возможность, что *в других веществах* поверхностные состояния подобным же образом продолжают из запрещенной зоны в разрешенную валентную зону (и, быть может, в зону проводимости). Такой случай, по-видимому, может иметь место в кристаллах типа $A^{IV}B^{VI}$ и $A^{III}B^V$, как будет показано в п. 3.

Дополнительные указания на то, что Ge имеет поверхность или вырожденного, или сильно вырожденного p -типа, были получены Про и Хандлером [224]. Они скалывали Ge в жидком N_2 . Хотя поверхность была покрыта жидким азотом (не была собственной), имело место, вероятно, только слабое взаимодействие ввиду инертности азота и низкой температуры. Тем не менее недавняя работа Хенслера [271], по-видимому, показала, что поверхность (111) Ge относится к невырожденному p -типу как после нагрева, так и без него. Распределение поверхностных состояний оказалось похожим на распределение, полученное для Si Алленом и Гобели [270], за исключением меньшего значения E_{sg} и близости E_1 к краю валентной зоны.

Некоторые теоретические расчеты в трехмерном случае, относящиеся к Ge, уже были рассмотрены ранее (в конце п. 1 § 17). Было показано, что их сопоставление с экспериментом в настоящее время не ведет к удовлетворительным результатам. Вероятно, это обусловлено перестройкой, которая не принималась во внимание.

3. Кристаллы типа $A^{III}B^V$ и $A^{IV}B^{VI}$

Отличительной чертой этих кристаллов по сравнению с Si и Ge является то, что связи в них частично ионные. Например, соединения ZnS и GaAs по оценкам имеют эффективные заряды $\pm 0,5e$ [156] и $0,2e$ [157] соответственно. Эти соединения встречаются в кристаллических модификациях цинковой обманки и вюрцита.

В них, подобно Si и Ge, имеются тетраэдрические связи. Отличие состоит в положении вторых ближайших соседей. Соединения BaO, CdO и др. со структурой каменной соли, по-видимому, еще не изучались с точки зрения их поверхностных свойств. Скол кристаллов со структурой цинковой обманки или вюрцита [197] осуществить легче, чем Si или Ge. Поэтому метод приготовления атомно чистой поверхности, используемый в настоящее время большинством исследователей, состоит просто в скалывании в сверхвысоком вакууме без нагрева. Некоторые поверхности скола [(110) цинковой обманки, (10 $\bar{1}$ 0) вюрцита], по-видимому, не перестраиваются (нет дополнительных рефлексов на дифракционной картине по методу ДЭНЭ), являются скомпенсированными (равное число ионов А и В в плоскости поверхности) и термически стабильными (не происходит изменений в картине рефлексов после нагрева). Исключением [231], вероятно, служит ZnO, игольчатые кристаллы которого могут быть сколоты с образованием полярных поверхностей {0001} [223, 283], по-видимому, неперестроенных [283]. Ниже будет показано, что ZnO также обладает особыми свойствами поверхностных состояний. Собственные полярные поверхности цинковой обманки {111} и вюрцита {0001} номинально состоят целиком либо из ионов А, либо из ионов В. Они могут быть получены с использованием техники бомбардировки ионами Ag и отжига¹⁾). Электростатический расчет показывает, что эти номинальные слои должны быть нестабильными [30]. Они, однако, могут быть стабилизированы [236] вакантными узлами, на долю которых приходится 1/4 узлов на поверхности (в согласии с интерпретацией Макрэм [233] данных по ДЭНЭ), или нагревом с образованием скомпенсированной микрофасеточной структуры поверхностей [236]. Экспериментально найдено, что нагрев поверхности {111} цинковой обманки приводит к образованию микрофасеточной структуры (110) [148], а нагрев поверхности {0001} вюрцита дает структуру (30 $\bar{3}$ 8) [230]. Главной трудностью в изучении

¹⁾ Для А- и В-компонент соединения АВ скорости очистки были, по-видимому, равны [284].

свойств поверхностных состояний этих полупроводников является невозможность контроля легирования в широких пределах из-за самокомпенсации. Например, CdS можно вырастить только *n*-типа, а ZnTe — только *p*-типа [285]. Как было сообщено, нагрев кристаллов CdS в сверхвысоком вакууме после скола приводит к неоднородности в концентрации доноров [286]. Поверхностные слои толщиной $d < 50 \text{ \AA}$ иногда имеют концентрацию доноров, приблизительно в 10 раз превышающую их концентрацию в объеме.

Данные по поверхностным состояниям у *скомпенсированных поверхностей* структур цинковой обманки (110) и вюрцита (11 $\bar{2}$ 0) были описаны для CdS, CdSe, CdTe, ZnO, ZnS, ZnSe и ZnTe Сванком [154], для InP и AlSb — Фишером [198] и [287], для InAs — Фишером с соавторами [288] и для GaAs, GaSb, InAs и InSb — Гобели и Алленом [197, 204]. Все эти результаты получены сходным путем с помощью измерения работы выхода ϕ_w и (или) порога фотоэмиссии для сколотых образцов при комнатной температуре. Во многих случаях была получена дополнительная информация из энергетического распределения фотоэмиссии или из экспериментов по фото-э. д. с., которые всегда выполнялись в одинаковых вакуумных условиях¹⁾.

Удалось ли систематизировать эти данные? Гобели и Аллен [197, 204] в 1965 г. предположили, что имеется следующая корреляция. Антимониды (GaSb и InSb) ведут себя подобно Ge, а арсениды (GaAs и InAs) аналогичны Si в том смысле, что у первой группы уровень Ферми закреплен у валентной зоны на поверхности или вблизи нее, тогда как в случае второй группы должна наблюдаться иная картина. Однако в 1967 г. более полные исследования Фишера и др. [288] показали, что в InAs уровень E_F закреплен у валентной зоны, что противоречит предшествующим результатам [197, 204]. Это показывает, что предложенная выше корреляция [197, 204] оказывается бесполезной.

В поисках новых зависимостей во время написания этого обзора было высказано предположение, что,

¹⁾ См. также работу [338]. — *Прим. перев.*

поскольку, как известно, рассматриваемые кристаллы являются частично ионными, может наблюдаться корреляция свойств с шириной запрещенной зоны в объеме, которая служит грубой мерой степени ионности [289—292]. Табл. 4 показывает, что на самом деле величина E_g коррелирует со многими характеристиками. Тринадцать соединений типа $A^{III}B^V$ и $A^{II}B^{VI}$, расположены в ней в порядке возрастания E_g : от 0,18 эВ для InSb до 3,6 эВ для ZnS. В последней колонке даны ссылки на экспериментальные работы. Закреплен ли уровень Ферми точно на краю валентной зоны или вблизи него, указано в третьей колонке словами *да* или *нет*. *Да* отмечает те кристаллы, у которых $E_g < 1$ эВ, а *нет* — те, у которых $E_g > 1$ эВ. (Для сравнения в таблицу вклю-

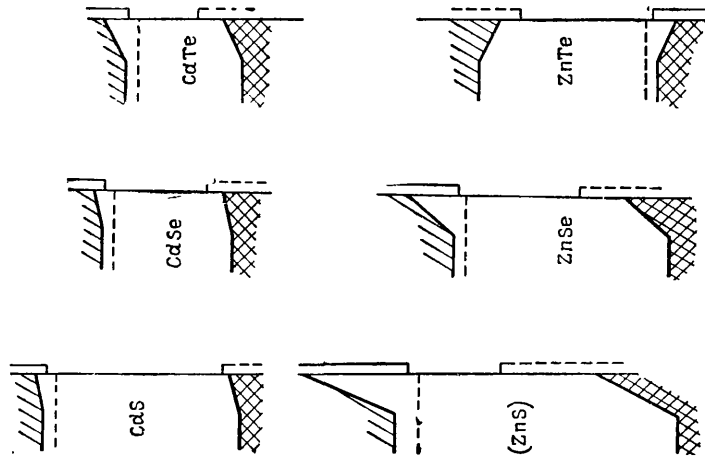
Таблица 4

**Характеристика изгиба зон в сколотых кристаллах
типа $A^{III}B^V$ и $A^{II}B^{VI}$**

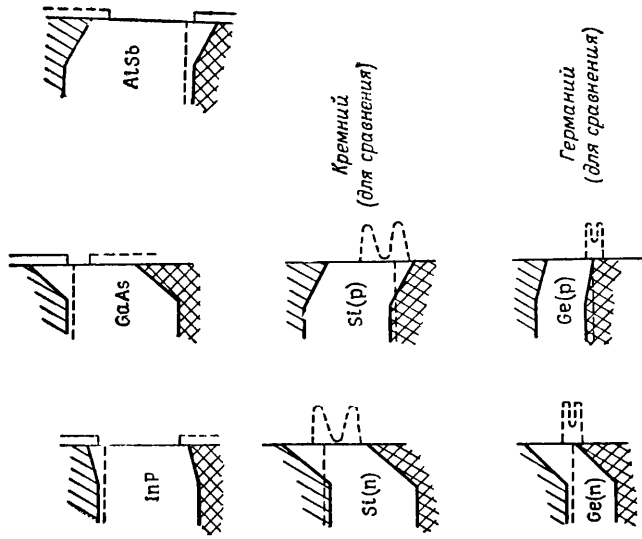
Кристалл	Энергетическая щель E_g , эВ	Закреплен ли уровень Ферми E_F у валентной зоны	Изгиб зон	Литературный источник
ZnS	3,6	нет	$n\uparrow$	[154]
ZnO	3,3	»	$n\uparrow$	[154]
ZnSe	2,7	»	$n\uparrow$	[154]
CdS	2,5	»	$n\uparrow$	[154]
ZnTe	2,2	»	$p\downarrow$	[154]
CdSe	1,7	»	$n\uparrow$	[154]
AlSb	1,6	»	$p\downarrow$	[287]
CdTe	1,5	»	$n\uparrow$	[154]
GaAs	1,4	»	$n\uparrow$	[207]
InP	1,3	»	$n\uparrow$	[198, 199]
Si	1,09	»	$n\uparrow p\downarrow$	[176]
GaSb	0,7	да		[204] *
Ge	0,67	»		[270]
InAs	0,35	»		[288]
InSb	0,18	»		[204]

* См. также работу [338]. — Прим. перев.

чены также Si и Ge.) Таким образом, возникает указанная корреляция. Ее можно проследить еще дальше, т. е. можно исследовать, будут ли зоны загнуты вверх или вниз, если кристалл легирован примесью n - или p -типа. Для случая $E_g > 1$ эВ результаты сведены в четвертой колонке табл. 4 и схематически показаны на фиг. 41 и 42. (Данные по Si и Ge, полученные Алленом и Гобели, также приведены для сравнения.) Обозначения $n\uparrow(p\downarrow)$ указывают, что кристаллы n -типа (p -типа) имеют зоны, прогнутые вверх (вниз) у поверхности. Из таблицы и фигур следует, что имеются только ситуации $n\uparrow$ и $p\downarrow$, в то время как $n\downarrow$ и $p\uparrow$ отсутствуют. Такое положение свидетельствует о том, что подобные кристаллы ведут себя так, как ожидалось в случае «ионных поверхностных состояний». Описанная ситуация показана на фиг. 34, δ и была рассмотрена ранее в § 1, 4 и 10. Джонс [269] теоретически пришел к такому же выводу, рассматривая ионное возмущение структуры поверхностных состояний в Si. Именно такое возмущение снимает вырождение вдоль направления $\bar{X}-\bar{M}$ на фиг. 39. Экспериментально собственные поверхностные состояния акцепторного типа, расположенные ниже зоны проводимости, вызывают изгиб зон вверх у поверхности кристалла n -типа. Аналогичным образом состояния донорного типа выше валентной зоны изгибают зоны вниз у поверхности кристалла p -типа. Правило $n\uparrow p\downarrow$ оказывается нарушенным, когда $E_g < 1$ эВ. При этом энергия E_F закреплена точно на краю или вблизи валентной зоны. Такого нарушения и на самом деле следовало ожидать, поскольку кристаллы с очень малыми запрещенными зонами вряд ли можно считать ионными. Таким образом, некоторые качественные черты собственных поверхностных состояний у скомпенсированных поверхностей соединений элементов III—V и II—VI групп легко можно объяснить, исходя из простых соображений, основанных на их ионной природе. Коутецкий и Томашек [293] рассчитали свойства поверхностных состояний полярных поверхностей (111) и $(\bar{1}\bar{1}\bar{1})$ цинковой обманки, которые, как известно, перестраиваются. Было использовано обобщение их более ранних моделей [26, 266], описывающих эффект ионности путем



Ф и г. 41. Схема искривления зон и собственных по-
верхностных состояний в некоторых кристаллах
типа $A^{III}B^{VI}$.



Ф и г. 42. Схема искривления зон и собственных по-
верхностных состояний в некоторых кристаллах
типа $A^{III}B^V$, а также в Si и Ge.

чередования вместе с А и В кулоновских (α) и резонансных (γ) интегралов. Левин и Фриман [135] вычислили закон дисперсии поверхностных состояний у поверхности (110) цинковой обманки при различных положениях поверхностных атомов типа А и В. В зависимости от положений поверхностных атомов и кулоновских энергий могут появиться зоны поверхностных состояний А- и В-типа.

Количественный расчет поверхностных состояний не был проведен еще ни в одном из этих кристаллов. Чтобы это осуществить, необходимо поступить следующим образом. Сначала определить V_D для всех указанных кристаллов и для каждого значения параметра легирования ξ . Связь параметра ξ с нескомпенсированной концентрацией N_D или N_A доноров или акцепторов получается из уравнения типа (11.2). Позднее был проведен анализ, который показал, что энергии поверхностных состояний слабо чувствительны к точному значению эффективной массы m^* или диэлектрической проницаемости K . Зависимость энергий от этих параметров выражается в форме $\sim kT \ln(m^*/K^2)^{1/4}$. В работах Фишера по InP [198, 199] и AlSb [287] величины V_D не были получены точно, однако их легко можно вычислить, зная концентрацию примеси и эффективную массу (необходимую для вычисления N_c или N_v). Говоря конкретно, для InP $N_D = 6 \cdot 10^{15} \text{ см}^{-2}$ и $m^* = 0,073m_0$ [294]. Отсюда $N_c = 4,8 \cdot 10^{17} \text{ см}^{-3}$ и $\xi = 0,11 \text{ эВ}$. Воспользовавшись фиг. 1 из работы Фишера, находим $V_D = 0,13 \text{ эВ}$. Аналогично для AlSb имеем $N_A = 2 \cdot 10^{17} \text{ см}^{-3}$ и $m^* = 0,9m_0$ [287] (эффективная масса по плотности состояний для дырок). Отсюда $N_v = 2,1 \cdot 10^{19} \text{ см}^{-3}$ и $\xi = 0,12 \text{ эВ}$. С помощью фиг. 2 из той же работы находим $eV_D = -0,24 \text{ эВ}$. Полученные значения $eV_D = 0,13$ и $0,24 \text{ эВ}$ малы (как это признается Фишером), но пренебречь ими нельзя. Используя значение $m^* \sim 0,2m_0$ для всех соединений типа $A^{II}B^{VI}$ (с точностью до множителя ~ 2), с помощью ξ можно вычислить положение примесных уровней (взяв $\xi = \Delta\phi_s - eV_D$ из табл. 1 работы Сванка [154]). Имеется некоторая несогласованность в указанных Сванком пределах концентрации для ZnTe.

Параметры поверхностных состояний в неко-

	E_g (эВ)	eV_D (эВ)	ζ (эВ)	$L \cdot 10^{-6}$ (см)
CdS	2,47	0,07	0,15	5,1
CdSe	1,67	0,12	0,15	5,1
CdTe	1,50	0,23	0,16	6,2
ZnO	3,25	0,01	0,10	1,9
(ZnS)	(3,6)	(1,2)	(0,30)	(0,01)
ZnSe	2,73	0,58	0,15	5,1
ZnTe	2,23	-0,18	0,15	5,1
GaAs	1,40	0,59	0,05	1,4
InP	1,29	0,13	0,11	4,7
AlSb	1,62	-0,24	0,12	0,83

В целях удобства в табл. 5 выписаны значения V_D и ζ для всех соединений с $E_g > 1$ эВ совместно с величинами запрещенных зон (во всех случаях измеренных при комнатной температуре). Соответствующие схемы изгиба зон показаны на фиг. 41 и 42. Ввиду недостатка дополнительной информации плотности электронных и дырочных поверхностных состояний (поверхностных ловушек) приняты равными, что и показано пунктирными линиями. Сюда не включена схема для соединения ZnO, так как для него характерны почти плоские зоны ($eV_D = 0,01$ эВ). Результаты для ZnS, согласно Сванку [154], сомнительны (поэтому это соединение в таблице заключено в скобки). Оценка более свежих данных для ZnTe¹⁾ дает $eV_D \sim -0,1$ эВ. На фиг. 42 для сравнения показаны также две схемы для легированных Si [176] и Ge [93]. Следующим шагом должно быть вычисление пространственного заряда Q_{sc} с помощью V_D и ζ . Зная Q_{sc} , можно указать положение поверхностных состояний. Поскольку считается, что обедненные слои у всех рассмотренных нами поверхностей тонкие, формулы оказы-

¹⁾ R. Swank, частное сообщение.

торых кристаллах типа AIII BV ($E_g > 1$ эВ)

Таблица 5

F_s	$Q_{sc}/e \cdot 10^{10}$ (см ⁻²)	m^*/m_0	$\Delta E_{1,2}$ (эВ)	E_{t1} (эВ)	E_{t2} (эВ)
2,0	5,4	0,2	0,090	0,13	—
2,8	7,6	0,2	0,082	0,20	—
4,1	9,0	0,2	0,077	0,31	—
0,4	3,0	0,2	0,102	0,01	—
(9,7)	(0,3)	(0,2)	(0,125)	(1,38)	—
6,6	18	0,2	0,060	0,67	—
3,5	-9,4	0,2	0,075	—	0,26
6,7	6,7	0,07	0,057	0,58	—
2,8	7,8	0,07	0,055	0,18	—
4,1	-65,0	0,9	0,067	—	0,29

ваются достаточно простыми. Для образцов p -типа имеем

$$Q_{sc} = -eN_A L F_s, \quad (17.1)$$

$$L = \left(\frac{KkT}{4\pi e^2 N_A} \right)^{1/2}, \quad (17.2)$$

$$F_s = \sqrt{2} [|v| + \exp(-|v|) - 1]^{1/2}, \quad (17.3)$$

$$v = \frac{eV_D}{kT}. \quad (17.4)$$

Для образцов n -типа аналогичные соотношения можно получить заменой N_A на N_D и переменной знака у Q_{sc} . Для всех веществ принято (с точностью до множителя 2), что $K \sim 10$. Значения дебаевского радиуса экранирования L , функций объемного заряда F_s и самого объемного заряда Q_{sc} для каждого из веществ приведены в табл. 5.

Из результатов § 12 следует, что распределение поверхностных состояний нельзя однозначно установить до тех пор, пока уровень Ферми и температура не меняются заметным образом. Тем не менее можно попытаться постулировать вид распределения, подобный принятому в гл. 6, а именно в ионных и частично ионных

кристаллах существуют зона состояний акцепторного типа ниже зоны проводимости и зона состояний донорного типа выше валентной зоны. Предполагается, что эти состояния имеют распределение *ионного типа* (фиг. 34, *д*). При этом ввиду недостатка информации глубины поверхностных электронных и дырочных ловушек принимаются равными. Это распределение аналогично представленному на фиг. 34, *б*, за исключением того, что зоны поверхностных состояний сливаются соответственно с валентной зоной и зоной проводимости. Будут ли поверхностные состояния, находящиеся *внутри валентной зоны и зоны проводимости*, локализованными или резонансными, зависит от пересечения дисперсионных кривых в *k*-пространстве, как это отмечалось в § 8. Однако этот вопрос в данном случае для нас не очень существен, поскольку легко доступны для измерений только те состояния, которые попадают внутрь объемной запрещенной зоны. Пусть E_2 (так же как на фиг. 34, *д*) означает верхнюю границу зоны донорных поверхностных состояний и пусть уровень Ферми E_F находится выше ее (поверхностные состояния не вырождены). Тогда в максвелловском приближении для двумерного электронного газа получаем

$$Q_{ss} = e\bar{N}kT \exp\left(-\frac{E_F - E_2}{kT}\right). \quad (17.5)$$

Здесь $\bar{N}$ определяется уравнением (12.7). Допустим, как и в гл. 6, что для веществ с большой запрещенной зоной эффективные массы m_s^* (поверхностных состояний) и m^* (объемных состояний, с которыми первые связаны) имеют сравнимую величину (скажем, с точностью до множителя ~ 2) [289—292, 135]¹⁾. Это может быть оправдано только для очень хорошо разделенных акцепторных и донорных поверхностных состояний [135], что, к счастью, как раз имеет место в случае веществ с большой запрещенной зоной $E_g > 1$ эВ (показано на фиг. 41 и 42). Таким образом, $\bar{N}$ однозначно определяется величиной m^* и в отсутствие внешних элек-

¹⁾ См. также S. Freeman, Phys. Rev., **B2**, 3272 (1970) и [135].

трических полей

$$Q_{ss} + Q_{sc} = 0. \quad (17.6)$$

Отсюда величину $E_F - E_2$ можно вычислить по формуле

$$E_F - E_2 = kT \ln \left(- \frac{Q_{sc}}{eNkT} \right). \quad (17.7)$$

Для кристаллов n -типа вместо $E_F - E_2$ берется $E_1 - E_F$. Можно в общем случае ввести параметры $\Delta E_1 = E_1 - E_F$ и $\Delta E_2 = E_F - E_2$, как показано на фиг. 34, δ . Меньший из этих параметров обозначим через ΔE ; значения его приведены в табл. 5. Как будет видно из дальнейшего, ΔE меняется с m^* и K , так как эти величины входят под знаком кубического и квадратного корня в логарифм. Эта слабая зависимость частично оправдывает многочисленные аппроксимации, сделанные ранее. Выражение (17.7) можно переписать в виде

$$\Delta E = \frac{1}{2} \zeta + kT \ln \frac{(8\pi e^2)^{1/2} (2\pi m^* kT)^{1/4}}{(hKF_s^2 kT)^{1/2}}. \quad (17.8)$$

Таким образом, $\Delta E \approx 1/2 \zeta$, если не принимать в расчет медленно меняющуюся величину, которая в интересующих нас случаях практически исчезает. Это подтверждается рассмотрением табл. 5. Также, поскольку $\Delta E \geq 2kT$, предположение о максвелловском распределении оправдано при комнатной температуре. Наконец, глубину ловушек относительно краев зон можно определить как

$$E_{t1} = E_c - E_1, \quad (17.9a)$$

$$E_{t2} = E_2 - E_v. \quad (17.9б)$$

Их значения приведены в последних колонках табл. 5.

Объединяя эти результаты, можно показать, что в первом приближении

$$E_{t1} \approx |eV_D| + \frac{1}{2} \zeta, \quad n\text{-тип}, \quad (17.9в)$$

$$E_{t2} \approx |eV_D| + \frac{1}{2} \zeta, \quad p\text{-тип}, \quad (17.9г)$$

где V_D — измеренное искривление зон. Для образцов n -типа (p -типа) зоны изогнуты вверх (вниз) у поверхности, как это показано на фиг. 41 и 42. Поскольку предполагается, что E_{t1} и E_{t2} не зависят от легирования, то из (17.9в) и (17.9г) следует, что чем сильнее легирование (меньше ξ), тем больше изгиб зон V_D . В настоящее время мы не располагаем данными, подтверждающими это предсказание. С другой стороны, если считать, что обе величины m_s^* и E_{t1} неизвестны, то для их определения требуется по крайней мере две группы данных (для различно легированных кристаллов). Указанные выше количественные результаты получились в предположении, что $m_s^* \approx m^*$ с точностью до множителя ~ 2 . Весьма полезным методом измерения энергий поверхностных состояний должны быть, вероятно, эксперименты с эффектом поля при постоянном токе, выполненные в сверхвысоком вакууме. Такие эксперименты возможны в случае соединений типа $A_{III}B^V$ и $A_{III}B^{VI}$ (но невозможны на Si и Ge) из-за малой величины поверхностного заряда: $Q_{ss}/e \approx 10^{10} - 10^{11} \text{ см}^{-2}$, как видно из табл. 5. Эту оценку можно получить, если просто положить $m^* = 0,2m_0$, $\Delta E \sim 0,1 \text{ эВ}$ и взять комнатную температуру. Тогда для Q_{ss}/e имеем из (17.5)

$$\frac{Q_{ss}}{e} \approx 4,2 \cdot 10^{14} (0,2) (0,025) e^{-4} \approx 4 \cdot 10^{10} \text{ см}^{-2}.$$

Эксперимент с эффектом поля при постоянном токе, выполненный в условиях, когда наступает вырождение зоны поверхностных состояний ($\Delta E = 0$, $E_F = E_t$), приводит к большей величине заряда:

$$\frac{Q_{ss}}{e} \approx 4,2 \cdot 10^{14} (0,2) (0,025) \approx 2 \cdot 10^{12} \text{ см}^{-2}.$$

Если же $E_t \sim 0,15 \text{ эВ}$, то условие плоской зоны дает предельное значение заряда:

$$\frac{Q_{ss}}{e} \approx 4,2 \cdot 10^{14} (0,2) (0,15) \approx 1 \cdot 10^{13} \text{ см}^{-2},$$

которое, по-видимому, трудно получить на практике.

Сравним теперь в деталях результаты по поверхностным состояниям, собранные в табл. 5 и на фиг. 41 и 42.

1) Для тщательно изученной группы соединений $\text{Cd} - \text{CdS}$, CdSe и CdTe глубина поверхностных ловушек E_{11} составила 0,13, 0,20 и 0,31 эВ соответственно. Заметна определенная тенденция: для более ионных кристаллов величина E_{11} меньше.

2) Группа соединений $\text{Zn} - \text{ZnO}$, ZnS , ZnSe труднее поддается описанию. Для ZnO $eV_D \sim 0,01$ эВ, что совершенно выпадает из всех известных данных по кристаллам типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$ и $\text{A}^{\text{II}}\text{B}^{\text{VI}}$. Указанное значение eV_D для ZnS , как теперь известно, слишком велико¹⁾. Соединение ZnTe относится к p -типу. Таким образом, в настоящее время не удается выявить закономерности в свойствах поверхностных состояний соединений Zn .

3) Глубина поверхностных ловушек (порядка нескольких десятых электрон-вольта) находится в согласии с численными оценками, основанными на методе Маделунга [30] (§ 10).

4) Можно предположить, что заряд Q_{ss} действительно связан с краевыми состояниями, отвечающими ступенькам и гребням рельефа, которые возникают в процессе скола. Однако Сванк [154] получил подобные результаты для искривления зон кристаллов с различной степенью совершенства скола. Фишер [295] также получил сходные результаты по фотоэмиссии для сколотого и изломанного GaAs . Хотя сейчас и нельзя окончательно исключить возможность краевых состояний, но уже моделей собственных поверхностных состояний более чем достаточно, чтобы объяснить результаты для различных кристаллов и концентраций примеси.

5) Исследователи, занимающиеся контактами металл — полупроводник, используют эмпирическое правило [296—298], которое гласит, что *контактные состояния у контакта металл — полупроводник сильнее выражены в тех кристаллах, где меньше ширина запрещенной зоны (или ниже степень ионности)*. Это правило с большими оговорками применимо и в случае собственных поверхностных состояний, как показано выше. Два

¹⁾ См. примечание на стр. 204.

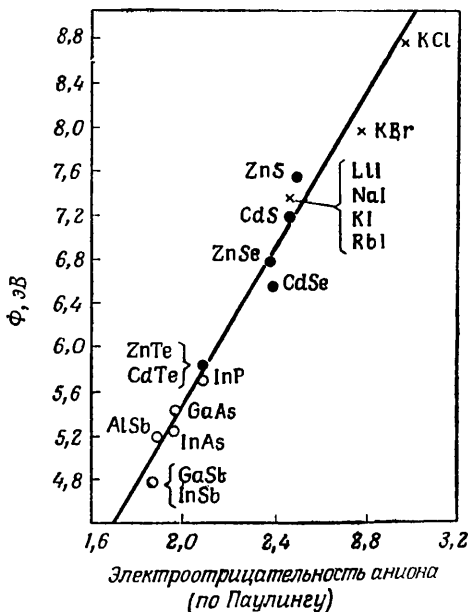
основных исключения — это соединение ZnO и ZnS , что подтвердили эксперименты, выполненные на нескольких образцах. Последние данные указывают, что значение E_i в ZnS много меньше, чем было сообщено ранее¹⁾. Тернер и Родерик [299] показали, что контактные состояния у контакта металл — диэлектрик *сильно* зависят от того, была ли поверхность атомно чистой перед напылением или нет.

6) Поверхностные состояния акцепторного типа у различных поверхностей кристалла CdS , по-видимому, похожи. Сванк [154] получил $eV_D = 0,07$ эВ для сколотой поверхности $(11\bar{2}0)$ CdS , а Кэмпбелл и Фарнsworth [230] нашли $eV_D = (0,02 - 0,06)$ эВ для бомбардированной и отожженной поверхности (0001) CdS (Cd -поверхность).

Рассмотрим теперь кратко те данные, на которых базируются эти выводы. Сванк [154] провел эксперименты, применяя скол в вакууме ($\sim 10^{-11}$ мм рт.ст.), и измерил потенциал Φ , экстраполируя к закону кубического корня. Также отмечалось наличие остаточного «хвоста», который мог быть обусловлен эмиссией либо с поверхностных состояний, либо из примесных состояний в объеме. К такому же выводу позднее пришел Фишер [198, 199]. Независимое доказательство того, что величина Φ (закон кубического корня) отвечает краю валентной зоны (как принимал Сванк), дала серия экспериментов с фото-э. д. с., выполненных параллельно. Измеренные с помощью фото-э. д. с. значения V_D для CdSe и CdTe составили 0,10 и 0,29 эВ соответственно. Это прекрасно согласуется с соответствующими значениями 0,12 и 0,23 эВ, полученными с помощью Φ и ϕ_w . По причине, указанной Сванком, CdS ведет себя аномально в экспериментах с фото-э. д. с. Если для определения Φ вместо закона кубического корня принять линейную зависимость, то значение V_D , полученное другими методами, будет отличаться более чем на 1 эВ. Таким образом, интерпретация Сванка представляется достаточно надежной. Было исследовано двенадцать образцов CdS , пять ZnSe , несколько образцов CdTe и ZnO

¹⁾ См. примечание на стр. 204.

и по одному CdSe, ZnS и ZnTe¹⁾. Сванк предложил интерпретировать данные с помощью полуклассической модели Маделунга [30]. Он сопоставил также значения Φ , полученные рядом авторов для соединений элементов III—V, II—VI и I—VII групп, с величиной электроотрицательности анионов. Результат показан на фиг. 43. Заметим, что через точки, отвечающие семнадцати соединениям, проходит совершенно прямая линия (ZnO вновь является исключением — он оказался за пределами графика). Этот результат показывает, что валентная зона всех этих соединений относится к анионному типу в согласии с моделями Маделунга [30], ЛКАО [33] и Матье [34]. Фишер и др. [288] развили аналогичную технику измерения, используя ϕ_w , Φ и распределение фотоэлектронов. Фишер предпочитает не применять кубической экстраполяции для порога, а предполагает для определения поверхностных свойств полупроводника [206] использовать измерения энергетического распределения. Гобели и Ален [197] дали прекрасный обзор фотоэмиссионных свойств, включающий данные вплоть до 1965 г. Ван-Лаар и Шир [186] исследовали потенциал Φ для GaAs разной степени леги-

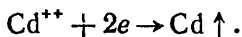


Фиг. 43. Порог фотоэмиссии Φ как функция электроотрицательности неметаллических компонент (из работы [154]).

¹⁾ См. примечание на стр. 204.

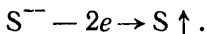
рования. Однако они не измеряли ϕ_w и, следовательно, не получили искривления зон при произвольном легировании. В то же время ими было отмечено, что уровень Ферми может сместиться примерно на 1 эВ ($E_g = 1,4$ эВ) при переходе от образца *n*-типа к образцу *p*-типа. Термины «стабилизированный» и «нестабилизованный» использовались ими, по-видимому, в применении к взаимно исключаящим группам взамен постепенного перехода, который подразумевается на фиг. 41 и 42. Было получено большое количество данных относительно величины Φ для самых различных веществ с одинаковой степенью легирования [205]. Эта информация, однако, бесполезна для определения поверхностных состояний, пока она не дополнена измерениями ϕ_w и (или) энергетического распределения электронов фотоэмиссии.

Дополнительное и независимое свидетельство в пользу существования поверхностных состояний ионного типа, показанных на фиг. 34, *д*, было получено Соморjai и Лестером [211]. Они изучали скорость испарения CdS при высоких температурах в глубоком вакууме и нашли, что эта скорость может возрасти примерно в 5 раз при одновременном облучении образца светом с частотой, большей, чем ширина запрещенной зоны. Поскольку со свободной поверхности непрерывно идет испарение и она обновляется, разумно считать такую поверхность чистой при исследовании собственных поверхностных состояний. Было найдено, что стадией, лимитирующей скорость испарения, является захват свободных электронов и дырок на поверхности. Используя предложенную выше модель ионных поверхностных состояний, эти данные можно объяснить следующим образом. В рамках простой классической ионной теории кристалл CdS состоит из ионов Cd^{++} и S^{--} . Для испарения нейтрального кадмия, который присутствует в газовой фазе, необходимо присоединить два электрона. Таким образом, реакция имеет вид



Как показано на фиг. 34, *д*, эти два электрона могут находиться в акцепторных поверхностных ловушках. При этом стадией, лимитирующей скорость испарения

Cd, будет захват двух электронов у поверхности в соответствии с экспериментом. Тот же аргумент применим в случае испарения серы, когда



Здесь потеря двух электронов эквивалентна присоединению двух дырок. Эти дырки могут находиться у поверхности в поверхностных ловушках донорного типа, как показано на фиг. 34, д. Согласно этим данным, при испарении CdS должны присутствовать состояния и донорного, и акцепторного типа.

В заключение следует упомянуть о полупроводниках типа $Al^{IV}V^{VI}$, таких, как PbS, PbSe, PbTe и т. д. Эти соединения применяются в качестве фотоэлементов в инфракрасной области. На *чистых* поверхностях было, видимо, выполнено только незначительное число работ. Оман [300] и Оман и Приоло [301] измерили ϕ_w и Φ для поверхности (100) PbS, сколотого в сверхвысоком вакууме. Полученные результаты по поверхностным состояниям не ясны ввиду низкого качества скола.

4. Несобственные поверхностные состояния

Для большей полноты желательно сравнить собственные поверхностные состояния с несобственными. Как уже отмечалось, большинство известных данных по поверхностным состояниям относится к несобственным состояниям, что объясняется техническими причинами. В этом коротком разделе под несобственными поверхностными состояниями понимаются просто все поверхностные состояния на свободной поверхности, не являющиеся собственными. Поверхности, отнесенные к типу несобственных, следующие:

1) Поверхности с адсорбатом неизвестной природы. Известно, что собственные поверхностные состояния могут быть пассивированы адсорбатами. Таков, например, CdS, обработанный на воздухе [302—304] и не очищенный в сверхвысоком вакууме с помощью аргоновой бомбардировки и отжига. Нагрев в сверхвысоком вакууме без бомбардировки аргоном редко очищает поверхность. Проблема нагрева особенно серьезна в случае бинар-

ных соединений, поскольку компоненты имеют различные давления паров.

2) Поверхности с известными адсорбированными (химически или физически) веществами. Например, свежесколотый в сверхвысоком вакууме Si, подвергнутый действию малого количества цезия [253] или кислорода [217]. Адсорбция может быть, а может и не быть обратимой. Аналогично адсорбция может давать и может не давать картину ДЭНЭ, отвечающую хорошо определенной двумерной поверхностной фазе.

3) Поверхности с одним или двумя монослоями адсорбатов. Здесь нет четкого различия между поверхностными и контактными состояниями, так как трехмерные фазы часто фрагментарны и плохо развиты. Примером может служить кристалл Si с несколькими монослоями кислорода [305].

4) Контакты между двумя твердыми веществами. Сюда относятся контакты полупроводник — окисел и полупроводник — металл.

5) Поверхность между жидкой и твердой фазами. Например, Ge, сколотый в чистом жидком азоте [224] или в другом жидком диэлектрике [225]. Другой пример — ZnO в контакте с электролитом [306].

Одной из общих черт несобственных поверхностей является то, что их физические, химические и структурные свойства *не удается* легко описать. Здесь обнаруживаются эффекты, которых нет у собственных поверхностных состояний. На результаты в общем случае влияют следующие факторы: 1) вещество адсорбата, давление, время экспозиции и температура поверхности; 2) предыстория обработки и роста; например, рост окисла из газовой фазы и сухой рост окисла из плазмы дают разные электрические свойства; 3) амплитуда и частота приложенного напряжения: в слоях окислов иногда имеются быстрые состояния, а иногда медленные, возможен также дрейф примесных ионов, и, наконец, 4) напряжения, возникающие в процессе роста у контакта двух твердых тел. Мы здесь не будем пытаться дать исчерпывающий обзор обширных сведений по несобственным поверхностным состояниям. Ниже будут рассмотрены только некоторые отдельные резуль-

таты, касающиеся быстрых состояний. По определению быстрые состояния находятся практически у интересующей нас поверхности или непосредственно вблизи нее.

К сожалению, некоторые экспериментаторы смешивают собственные и несобственные поверхностные состояния. Например, Мейни с соавторами [302—304] делали это в отношении CdS, а Хафф и др. [29] — в отношении InSb. В обоих случаях экспериментальные данные относились к поверхностям, приготовленным разрезанием, полировкой и травлением либо ростом из газовой фазы. Эти поверхности не были ни сколоты в сверхвысоком вакууме, ни очищены аргоновой бомбардировкой и отжигом. В случае CdS [302—304] в контакте с поверхностью находилась майларовая или слюдяная прослойки. Эти же авторы предполагали, что исследуемые ими поверхностные состояния — собственные, и сравнивали свои экспериментальные результаты с теориями собственных поверхностных состояний. Полученные ими таким образом распределения поверхностных состояний оказались в обоих случаях сильно отличными от наблюдавшихся ранее для собственных поверхностей CdS [154] и InSb [204]. Точнее говоря, Мейни и др. [302—304] не получили поверхностных состояний на поверхности $(11\bar{2}0)$ и полярных поверхностях несобственного полупроводника CdS. Напротив, Сванк [154] обнаружил собственные поверхностные состояния на поверхности $(11\bar{2}0)$ CdS, а Кэмпбелл и Фарнсворс [230] нашли собственные поверхностные состояния у поверхности (0001) CdS, как было упомянуто в предыдущем разделе. Аналогично Хафф и др. [29] получили распределение поверхностных состояний в InSb (110) , подобно фиг. 34, *е*, с нейтральным уровнем вблизи центра запрещенной зоны и большой постоянной плотностью состояний, покрывающей всю запрещенную зону. С другой стороны, Аллен и Гобели [204] нашли, что уровень Ферми закреплен около валентной зоны собственного InSb (110) . Примеры CdS и InSb были приведены здесь только для того, чтобы показать, что отсутствие четкого разграничения *собственных* и *несобственных* поверхностных состояний легко может привести к существенным недоразумениям.

Значительные усилия были предприняты с целью изучения контактных состояний $\text{Si} - \text{SiO}_2$. В настоящее время представляется, что этот контакт при помощи тщательной обработки может быть сделан почти идеальным в том смысле, что поверхностные состояния внутри *всей* запрещенной зоны $E_v < E_F < E_c$ не могут быть выявлены экспериментально [214, 307]. *Верхний* предел плотности при этом ($N < 10^{10} \text{ см}^{-2}$) получается с помощью метода «металл — окисел — полупроводник». Такая степень совершенства достигается весьма тщательным удалением H_2O и следов металлических примесей в выращенном слое окисла, а также надлежащей процедурой отжига. Наблюдавшиеся *собственные* поверхностные состояния в количестве порядка единицы на поверхностный атом Si (п. 1 § 17), по-видимому, исчезают после соответствующей «пассивации» окислением.

Контактные состояния у контакта металл — полупроводник изучались столь же интенсивно [296—298]. Оказалось, что существуют некоторые качественные соотношения между собственными поверхностными состояниями и контактными, даже несмотря на то, что окружение поверхностных атомов существенно изменилось. Кристаллы с большим числом собственных поверхностных состояний *внутри запрещенной зоны*, по-видимому, имеют большее число указанных контактных состояний. Этот эффект был выявлен при наблюдении изменения высоты барьера металл — полупроводник с изменением работы выхода металла. Для подтверждения этого грубого эмпирического правила необходимы дополнительные данные. Хейне [308] предложил теорию подобных контактов, которая принимает во внимание проникновение волновой функции металла в полупроводник.

Несобственные поверхностные состояния (быстрые состояния), как правило, укладываются в привычную схему. В частности, поверхностные состояния часто возникают парами, причем акцепторные состояния обычно лежат *выше* донорных, как на фиг. 34, *а*, *б*, *д* и *е*. Это приводит к изгибу зон вверх (вниз) у образцов *n*-типа (*p*-типа), подобно поверхности на фиг. 41 и 42. Иногда эти состояния считаются или дискретными (фиг. 34, *а*), или зонными (фиг. 34, *д*), или покрывающими всю основ-

ную запрещенную зону (фиг. 34, в), или как комбинация тех и других (фиг. 34, е). К указанной схеме были отнесены несобственные поверхностные состояния в Si [214, 309], Ge [224]¹⁾, AgCl [169], InSb [29, 310, 311], CdS [215, 312, 313] и GaAs [314, 315].

Эксперименты, выполненные на GaP [316] и GaAs [186], не привели ни к какой ясной схеме. Адсорбционные состояния четко проявились у ZnO [213, 223]. Исключение, вероятно, составили лишь поверхностные состояния донорного типа, обнаруженные в контакте Si—SiO₂ [307], которые иногда могут быть расположены выше края зоны проводимости Si.

Наблюдения, проведенные над несобственными поверхностными состояниями, приводят подчас к выводам, справедливым и для собственных поверхностных состояний. Например, Марк [215, 312] изучал фотопроводимость и свойства захвата пластинок (11 $\bar{2}$ 0) CdS, определяемые преимущественно поверхностью. Последние были воспроизводимо «очищены» обдуванием чистым азотом под давлением в 1 атм. Как уже отмечалось ранее, такой способ очистки не обеспечивает собственной поверхности. Говоря задним числом, это были, вероятно, несобственные поверхности. В результате своих экспериментов Марк впервые определенно заявил о существовании пар поверхностных состояний в CdS. Эта концепция «пар» создала основу для нового класса теорий. Как показано выше, эти теории, по-видимому, справедливы для большинства собственных поверхностных состояний в ионных и частично ионных кристаллах.

Теоретические модели несобственных поверхностных состояний очень многообразны. Они включают поля, связанные с точечными дефектами, хемосорбцию, физическую адсорбцию, катализ и т. д. Мы не будем пытаться здесь их рассматривать. Кроме того, как уже отмечалось, соответствующие экспериментальные данные зачастую недостаточно четко определены для их осмысленного сравнения с теорией.

¹⁾ Данные до 1967 г. по несобственным поверхностным состояниям в Si и Ge имеются в работах [7] и [8].

Глава 8

ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Авторы настоящего обзора преследовали в основном следующие цели: а) решить несколько различных задач о поверхности, используя широкий набор вычислительных методов, дав таким образом полезное компактное руководство для тех, кто занимается проблемами поверхности; б) исследовать поверхностные состояния на основе различных, дополняющих друг друга подходов; в) дать возможность неспециалистам познакомиться с основными представлениями и математическими методами с минимальной затратой усилий; г) суммировать экспериментальные методы, используемые при определении свойств поверхностных состояний, и д) собрать и рассмотреть данные, полученные для ряда различных кристаллов, и сравнить эти результаты с теорией.

Экспериментально было найдено, что, например, для Si и Ge все распределения поверхностных состояний оказываются парами донорно-акцепторных зон. В то же время трехмерные теоретические расчеты предсказывают только одну зону. Быть может, этот результат не удивителен, поскольку в теоретических расчетах еще не учитывались эффекты перестройки поверхности. Напротив, некоторые поверхности соединений типа $A^{III}B^V$ и $A^{II}B^{VI}$ не перестроены, и сравнение теории с экспериментом удовлетворительно. В частности, зоны поверхностных состояний акцепторного (донорного) типа возникают на удалении в несколько десятых долей электрон-вольта ниже (выше) зоны проводимости (валентной зоны). Зоны объемных и поверхностных состояний заметно перекрываются, и их эффективные массы оказываются примерно одинаковыми. Опыты с кристаллами, обладающими большой запрещенной зоной ($E_g > 1$ эВ), показывают, что образцы *n*-типа (*p*-типа) имеют зоны, загнутые вверх (вниз) у поверхности. Это согласуется с общей теоретической картиной ионных поверхностных состояний.

Среди некоторых электротехнических применений, в которых поверхностные состояния играют существенную роль, можно назвать следующие: диоды с барьером Шоттки, в которых «работает» истощенный слой полупроводника, покрытия передающей телевизионной трубки (видикона), «сэндвичи» типа металл — окисел — полупроводник, транзисторные системы, электрофотографические покрытия и фотокатоды с отрицательным сродством к электрону. В этом последнем случае используются полупроводники p -типа, у которых работа выхода меньше ширины запрещенной зоны. Здесь удастся в сотни раз увеличить выход фотоэмиссии в инфракрасной области спектра. Для получения таких низких значений работы выхода в случае Si (ширина запрещенной зоны равна 1,1 эВ) необходимо строго выдержать последовательность адсорбционных слоев цезия и кислорода [350—352]. Было показано, что собственные поверхностные состояния в Si исчезают в процессе адсорбции цезия, и вместо них образуется новая зона «цезиевых» поверхностных состояний с энергией примерно на 0,4 эВ ближе к зоне проводимости. Интерпретация этих несобственных поверхностных состояний довольно проста: им отвечают участки кремниевого и (или) кислородного субстрата, на которых атомы цезия оставляют часть своих зарядов, образуя поверхностный диполь с большим моментом. Определение длины диполя в системе Si/Cs/O с помощью геометрического метода дало значение 2,9 Å. Это находится в согласии со значением 2,8 Å, полученным расчетным путем, а также с помощью изучения поверхностных плазменных возбуждений [350—352].

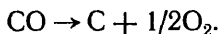
До недавнего времени оставался невыясненным вопрос, участвует ли в каталитических процессах собственно поверхность полупроводника или же главную роль играют особые центры, такие, как ступеньки, точечные дефекты или адсорбированные молекулы. Гипотеза активных центров не могла быть доказана на примере обычных каталитических систем, содержащих тонкий порошок полупроводника. Недавно, однако, были выполнены эксперименты с большими поверхностями монокристаллов ZnO, CdS и ZnS с использованием методов ДЭНЭ и оже-спектроскопии. Выяснилось следующее:

1) Подвергнутые ионной бомбардировке и отжигу поверхности (0001); (000 $\bar{1}$); (11 $\bar{2}$ 0) и (10 $\bar{1}$ 0) ZnO не могут быть заряжены потоком ионов [353] (отсутствуют накапливающие заряд несобственные поверхностные состояния). Однако с введением дефектов появляются несобственные поверхностные состояния, и кристалл заряжается. Дефекты могут быть созданы при желании действием HCl в виде пара или в процессе травления. Замечательной особенностью экспериментов с ZnO явилось наличие изгиба зон величиной примерно 50 эВ, причем это неравновесное состояние может сохраняться в темноте при комнатной температуре в течение десятков минут. Изучался процесс разрядки такой поверхности, названный спектроскопией фоторазрядки, при облучении светом с частотой, меньшей, чем ширина запрещенной зоны.

С помощью этой методики удалось идентифицировать энергетические уровни несобственных поверхностных состояний для каждой из упомянутых кристаллических поверхностей.

2) Относительно совершенная поверхность (11 $\bar{2}$ 0) CdS не присоединяет кислород; это показывают эксперименты по фотопроводимости. Однако адсорбция кислорода сильно нарастает при внесении поверхностных дефектов [354, 355].

3) Идеальный кристалл ZnS с поверхностью (110) не является катализатором реакции разложения CO:



Однако аналогичные поверхности, вырезанные под косым углом к плоскости (110), содержали далеко отстоящие друг от друга поверхностные ступеньки высотой порядка атомных размеров. Было показано [356], что эти ступеньки ответственны за катализ указанной реакции. Независимое доказательство того, что ступеньки атомных размеров играют важную роль в катализе, было недавно получено для металлических поверхностей [357].

В данном обзоре мы не пытались рассмотреть динамические аспекты, такие, как явления переноса, многочастичные эффекты, оптические эффекты и взаимодействие с поверхностными фононами, магнонами и

экситонами. В настоящее время поняты отчасти только статические свойства.

Дальнейшая тенденция развития теории поверхностных состояний может включать более интенсивное использование вычислительных машин, применение более реалистичных и изошренных моделей, способных описывать указанные выше динамические эффекты, и более точное рассмотрение поверхностной области и перестройки, если она имеется. В отношении эксперимента программа на будущее может содержать исследование кристаллов, концентрация примеси в которых меняется в широких пределах измерения при низких температурах с целью нахождения распределения поверхностных состояний, а также систематическое изучение новых классов веществ.

ЛИТЕРАТУРА¹⁾

1. *Grimley T. B.*, Advan. Catalysis., **12**, 1 (1960).
2. *Koutecký J.*, Advan. Chem. Phys., **9**, 85 (1965).
3. *Mark P.*, Catalysis Rev., **1**, 165 (1967).
4. *Davison S. G.*, Intern. Journ. Quantum. Chem., **2S**, 291 (1968).
5. *Levine J. D., Mark P.*, Phys. Rev., **182**, 926 (1969).
6. *Davison S. G., Stešlicka M.*, Quantum Surface Physics, Oxford, 1971.
7. *Many A., Goldstein Y., Grover N. B.*, Semiconductor Surfaces, Amsterdam, 1965.
8. *Frankl D. R.*, Electrical Properties of Semiconductor Surfaces, Oxford, 1967.
9. *Волькенштейн Ф. Ф.*, Электронная теория катализа на полупроводниках, Физматгиз, 1960.
10. *Born M., Oppenheimer J. R.*, Ann. Physik., **84**, 457 (1927).
11. *Davison S. G.*, в книге Physical Chemistry, vol. 3, ch. 2, New York, 1969.
12. *Löwdin P. O.*, Advan. Chem. Phys., **2**, 207 (1959).
13. *Kronig R., Penney W. G.*, Proc. Roy. Soc., **A130**, 499 (1931).
14. *Тамм И. В.*, Zs. Physik, **76**, 849 (1932); Phys. Z. Sowjet., **1**, 733 (1932).
15. *Maue A. W.*, Zs. Physik, **94**, 717 (1935).
16. *Goodwin E. T.*, Proc. Cambridge. Phil. Soc., **35**, 205, 221, 232 (1939).
17. *Shockley W.*, Phys. Rev., **56**, 317 (1939).
18. *Bardeen J.*, Phys. Rev., **71**, 717 (1947).
19. *Meyerhof W. E.*, Phys. Rev., **71**, 727 (1947).
20. *Shockley W., Pearson G.*, Phys. Rev., **74**, 232 (1948).
21. *Statz H.*, Zs. Naturforsch., **5a**, 534 (1950).
22. *Artmann K.*, Zs. Physik, **131**, 244 (1952).
23. *Hoffmann T. A.*, Acta Phys. Acad. Sci. Hung., **2**, 195 (1952).
24. *Lippmann B. A.*, Ann. Phys., **2**, 16 (1957).
25. *Koutecký J.*, Phys. Rev., **108**, 13 (1957).
26. *Koutecký J., Tomášek M.*, Phys. Rev., **120**, 1212 (1960).
27. *Antončík E.*, Journ. Phys. Chem. Solids, **21**, 137 (1961).
28. *Gatos H. C., Lavine M. C.*, Journ. Electrochem. Soc., **107**, 427, 433 (1960).
29. *Huff H. R., Kawaji S., Gatos H. C.*, Surface Sci., **10**, 232 (1968).
30. *Levine J. D., Mark P.*, Phys. Rev., **144**, 751 (1966).
31. *Davison S. G., Koutecký J.*, Proc. Phys. Soc., **89**, 237 (1966).
32. *Koutecký J., Davison S. G.*, Intern. Journ. Quant. Chem., **2**, 73 (1968).
33. *Levine J. D., Davison S. G.*, Phys. Rev., **174**, 911 (1968).
34. *Levine J. D.*, Phys. Rev., **171**, 701 (1968).
35. *Davison S. G., Cheng Y. C.*, Phys. Lett., **27A**, 592 (1968).

¹⁾ Литература, отмеченная одной звездочкой, добавлена переводчиком, двумя звездочками — авторами в корректуре. — *Прим. ред.*

36. Bloch F., Zs. Physik, **52**, 555 (1928).
37. Ziman J. M., Principles of the Theory of Solids, London and New York, 1965. (Имеется перевод: Дж. Займан, Принципы теории твердого тела, изд-во «Мир», 1966.)
38. Sachs M., Solid State Theory, New York, 1963.
39. Smith R. A., Wave Mechanics of Cristalline Solids, London, 1967.
40. Dirac P. A. M., Quantum Mechanics, London and New York, 1959. (Имеется перевод: П. Дирак, Принципы квантовой механики, Физматгиз, 1960.)
41. Streitwieser A., Jr., Molecular Orbital Theory for Organic Chemists, ch. 2, New York, 1966.
42. Salem L., The Molecular Orbital Theory of Conjugated Systems, ch. 1, New York, 1966.
43. Milne-Thomson L. M., The Calculus of Finite Differences, New York, 1951, p. 387.
44. Brand L., Differential and Difference Equations, ch. 9, New York, 1966.
45. Amos A. T., Davison S. G., Physica, **30**, 905 (1964).
46. Hoffmann T. A., Acta Phys. Acad. Sci. Hung., **1**, 175 (1951).
47. Aerts E., Physica, **26**, 1057 (1960).
48. Koutecký J., Trans. Faraday Soc., **54**, 1038 (1958).
49. Grimley T. B., Proc. Phys. Soc., **72**, 103 (1958).
50. Rymer T. B., Nuovo Cimento Suppl., **6**, Ser. X, 1, 294 (1957).
51. Жданов Г. С., Физика твердого тела, изд-во МГУ, 1961.
52. Lander J. J., Gobell G. W., Morrison J., Journ. Appl. Phys., **34**, 2298 (1963).
53. Marsh J. B., Farnsworth H. E., Surface Sci., **1**, 3 (1964).
54. Davison S. G., Physica, **37**, 539 (1967); Surface Sci., **10**, 369 (1968).
55. Stęślicka M., Acta Phys. Polon., **33**, 981 (1968); **34**, 875 (1968).
56. Phariseau P., Physica, **26**, 1192 (1960).
57. Stęślicka M., Wojciechowski K. F., Physica, **32**, 1274 (1966).
58. Stęślicka M., Acta Phys. Polon., **30**, 883 (1966).
59. Davison S. G., Stęślicka M., Int. Journ. Quantum Chem., **4**, 455 (1971).
60. Henzler M., Surface Sci., **9**, 31 (1968).
61. Hoffmann T. A., Acta Phys. Acad. Sci. Hung., **2**, 107 (1952).
62. Davison S. G., Ling T. Y., Ghosh U. S., Journ. Phys. Chem. Solids, **28**, 1921 (1967).
63. Arfken G., Mathematical Methods for Physicists, 2nd ed., New York, 1970. (Имеется перевод: Г. Арфкен, Математические методы в физике, Атомиздат, 1970.)
64. Davison S. G., Amos A. T., Journ. Chem. Phys., **43**, 2223 (1965).
65. Луфшиц Н. М., ЖЭТФ, **17**, 1017, 1076 (1947).
66. Baldock G. R., Proc. Cambridge Phil. Soc., **48**, 457 (1952).
67. Koster G. F., Slater J. C., Phys. Rev., **95**, 1167 (1954).
68. Löwdin P.-O., Journ. Math. Phys., **3**, 969 (1962).
69. Messiah A., Quantum Mechanics, vol. 1, p. 260; vol. 2, p. 712, Amsterdam, 1962.
70. Löwdin P.-O., Journ. Chem. Phys., **19**, 1396 (1951); Journ. Mol. Spectry., **10**, 12 (1963); **14**, 119 (1964).
71. Levine J. D., Amos A. T., Phys. Stat. Sol., **19**, 587 (1967).

72. Wannier G. H., Phys. Rev., **52**, 191 (1937).
73. Wannier G. H., Elements of Solid State Theory, London and New York, 1959, p. 173.
74. Baldock G. R., Proc. Phys. Soc., **66**, 1 (1953).
75. Tomášek M., Koutecký J., Czech., Journ. Phys., **B10**, 268 (1960).
76. Koutecký J., Czech. Journ. Phys., **B11**, 565 (1961).
77. Davison S. G., Derco G. L., Surface Sci., **6**, 323 (1967).
78. Fairbairn W. M., Surface Sci., **9**, 439 (1968).
79. Davison S. G., Grindlar J., Surface Sci., **11**, 99 (1968).
80. Feuer P., Phys. Rev., **88**, 92 (1952).
81. Fukushima M., Sci. Rept., **16**, 1 (1967).
82. Alstrup I., Johansen K., Phys. Stat. Sol., **28**, 555 (1968).
83. Davison S. G., Taylor N. F., Chem. Phys. Lett., **3**, 424 (1969).
84. Saxon D. S., Hutner R. A., Philips Res. Rep., **4**, 81 (1949).
85. Brillouin L., Journ. Phys. Radium., **1**, 377 (1930).
86. Morse P. M., Phys. Rev., **35**, 1310 (1930).
87. Slater J. C., Phys. Rev., **87**, 807 (1952).
88. Brillouin L., Wave Propagation in Periodic Structures, ch. 8, New York, 1953. (Имеется перевод: Л. Бриллюэн, М. Пароди, Распространение волн в периодических структурах, ИЛ, 1959.)
89. Mathieu E., Journ. Math. Pures Appl., **13**, 137 (1968).
90. Whittaker E. T., Watson G. N., Modern Analysis, ch. 19, London and New York, 1958. (Имеется перевод: Э. Т. Уиттекер, Дж. Н. Ватсон, Курс современного анализа, Физматгиз, 1963.)
91. McLachlan N. W., Theory and Applications of Mathieu Functions, New York, 1964, p. 99.
92. Johnson L. E., Conklin J. B., Jr., Pratt G. W., Jr., Phys. Rev. Lett., **11**, 538 (1963).
93. Herman F., Kuglin C. D., Cuff D. F., Kortum R. L., Phys. Rev. Lett., **11**, 541 (1963).
94. Conklin J. B., Jr., Johnson L. E., Pratt G. W., Jr., Phys. Rev., **137**, A1282 (1965).
95. Soven P., Phys. Rev., **137**, A1706 (1965).
96. Chow P. C., Liu L., Phys. Rev., **140**, A1817 (1965).
97. Loucks T. L., Phys. Rev. Lett., **14**, 693, 1072 (1965); Phys. Rev., **139**, A1333 (1965); **143**, 506 (1966).
98. Augmented Plane Wave Method, New York, 1967.
99. Foldy L. L., Wouthuysen S. A., Phys. Rev., **78**, 29 (1950).
100. Wood R. D., Callaway J., Bull. Am. Phys. Soc., **2**, 18 (1957).
101. Glasser M. L., Davison S. G., Intern. J. Quantum Chem., **39**, 867 (1969).
102. Schiff L. I., Quantum Mechanics, New York, 1955. (Имеется перевод: Л. Шифф, Квантовая механика, ИЛ, 1957.)
103. Śteślicka M., Davison S. G., Brown A. G., Proc. NBS Mater. Res. Symp., 3rd (Spec. Publ. 323), 1969, p. 477.
104. Śteślicka M., Davison S. G., Phys. Rev., **B1**, 1858 (1970).
105. Merzbacher E., Quantum Mechanics, New York, 1964, p. 99.
106. Jones H., The Theory of Brillouin Zones and Electronic States in Crystals, Amsterdam, 1962. (Имеется перевод: Г. Джонс, Теория зон Бриллюэна и электронные состояния в кристаллах, изд-во «Мир», 1968.)

107. *Turnbull H. W.*, Theory of Equations, Edinburgh, 1957.
108. *Davison S. G.*, *Stęślicka M.*, Journ. Phys. C. (Proc. Phys. Soc.), 2, 1802 (1969).
109. *Stęślicka M.*, Physica, 45, 506 (1970).
110. *Morse P. M.*, *Feshbach H.*, Methods of Theoretical Physics, New York, 1953. (Имеется перевод: Ф. Морс., Г. Фешбах, Методы теоретической физики, ИЛ, 1958—1960.)
111. *Aerts E.*, Physica, 26, 1047 (1960); см. также в книге Solid State Physics in Electronics and Telecommunications, vol. 1, part 1, New York, 1960, p. 628.
112. *Phariseau P.*, Physica, 26, 737 (1960).
113. *Aerts E.*, Physica, 26, 1163 (1960).
114. *Ghosh S. K.*, *Laskar A. K.*, Ind. Journ. Phys., 37, 534 (1963).
115. *Phariseau P.*, Physica, 27, 590 (1961).
116. *Phariseau P.*, Physica, 26, 1185 (1960).
117. *Phariseau P.*, Physica, 30, 608 (1964).
118. *Koutecký J.*, Phys. Stat. Sol., 1, 554 (1961).
119. *Jahnke E.*, *Emde F.*, Table of Functions with Formulae and Curves, New York, 1945. (Имеется перевод: Е. Янке, Ф. Эмде, Ф. Леш, Специальные функции. Формулы, графики, таблицы, изд-во «Наука», 1968.)
120. *Cohen M. L.*, *Bergstresser T. K.*, Phys. Rev., 141, 789 (1966).
121. *Kleinman I.*, *Phillips J. C.*, Phys. Rev., 125, 809 (1962).
122. *Mark P.*, в сб. Clean Surfaces: Their Preparation and Characterization for Interfacial Studies, New York, 1969.
123. *Mourad P.*, Phys. Stat. Sol., 12, 817 (1965).
124. *Heine V.*, Proc. Phys. Soc., 81, 300 (1963).
125. *Mott N. F.*, *Jones H.*, The Theory of the Properties of Metals and Alloys, New York, 1958.
126. *Heine V.*, в сб. Solid Surfaces, Amsterdam, 1964.
127. *Boudreaux D. S.*, *Heine V.*, Surface Sci., 7, 426 (1967).
128. *Blout E. I.*, Solid State Phys., 13, 305 (1962).
129. *Jones R. O.*, Proc. Phys. Soc., 89, 443 (1966).
130. *Stern R. M.*, *Perry J. J.*, *Boudreaux D. S.*, Rev. Mod. Phys., 41, 275 (1969).
131. *Phariseau P.*, Phys. Stat. Sol., 3, 2391 (1963).
132. *Chaves C. M.*, *Majlis N.*, *Cardona M.*, Solid State Commun., 4, 271 (1966).
133. *Pugh D.*, *Phys-Roberts C.*, *Tregold R. H.*, Proc. Phys. Soc., 80, 534 (1962).
134. *Jones R. O.*, Phys. Rev. Lett., 20, 992 (1968).
135. *Levine J. D.*, *Freeman S.*, Bull. Am. Phys. Soc., 14, 787 (1969); Phys. Rev., B2, 3255 (1970).
136. *Handler P.*, в сб. Semiconductor Surface Physics, Pennsylvania, 1957, p. 23.
137. *Allen F. G.*, *Fowler A. B.*, Journ. Phys. Chem. Sol., 3, 107 (1957).
138. *Handler P.*, Journ. Phys. Chem. Sol., 14, 1 (1960).
139. *Gatos H. C.*, *Lavine M. C.*, Journ. Phys. Chem. Sol., 14, 169 (1960).
140. *Gatos H. C.*, Journ. Appl. Phys., 32, 1232 (1961).
141. *Mariano A. N.*, *Hanneman R. E.*, Journ. Appl. Phys., 34, 384 (1962).

142. *Finn M. C., Gatos H. C.*, Surface Sci., **1**, 361 (1964).
143. *Taloni A., Haneman D.*, Surface Sci., **8**, 323 (1967).
144. *Sangster R. G.*, в сб. Compound Semiconductors, vol. 1, New York, 1967, p. 241.
145. *Haneman D.*, Journ. Phys. Chem. Sol., **14**, 162 (1960).
146. *Haneman D.*, Phys. Rev., **121**, 1093 (1961).
147. *Holt D. B.*, Journ. Appl. Phys., **31**, 2231 (1960).
148. *MacRae A. U., Gobeli G. W.*, в сб. Semiconductors and Semimetals, vol. 2, New York, 1966.
149. *MacRae A. U.*, Surface Sci., **4**, 247 (1966).
150. *Wells A. F.*, Structural Inorganic Chemistry, London, and New York, 1962.
151. *Seitz F.*, Modern Theory of Solids, New York, 1940, p. 76.
152. *Mott N. F., Gurney R. A.*, Electronic Processes in Ionic Crystals, New York, 1964, p. 86. (Имеется перевод: Н. Ф. Мотт, Р. Герни, Электронные процессы в ионных кристаллах, ИЛ, 1950.)
153. *Mott N. F.*, Proc. Phys. Soc., **49**, 258 (1937).
154. *Swank R. K.*, Phys. Rev., **153**, 844 (1967).
155. *Pauling L.*, The Nature of the Chemical Bonds, New York, 1960, p. 508. (Имеется перевод 1-го изд.: Л. К. Паулинг, Природа химической связи, Госхимиздат, 1947.)
156. *Birman J. L.*, Phys. Rev., **109**, 810 (1958).
157. *Wolff G. A.*, в книге Compound Semiconductors, vol. 1, New York, 1967.
158. *Mark P.*, Journ. Phys. Chem. Sol., **29**, 689 (1968).
159. *Ritche R. H.*, Surface Sci., **4**, 497 (1965).
160. *Raether H.*, Surface Sci., **8**, 233 (1967).
161. *Zuehlke R. W.*, Journ. Chem. Phys., **45**, 411 (1966).
162. *Geus J. W.*, Surface Sci., **2**, 48 (1964).
163. *Juretschke H. J.*, Surface Sci., **5**, 111 (1966).
164. *Hagstrum H. D., Takeishi Y., Becker G. E., Pretzer D. D.*, Surface Sci., **2**, 26 (1964).
165. *Knor Z., Müller E. W.*, Surface Sci., **10**, 21 (1968).
166. *Tsong T. T.*, Surface Sci., **10**, 303 (1968).
167. *Levine J. D., Gyftopoulous E. P.*, Surface Sci., **1**, 171 (1964).
168. *Gyftopoulous E. P., Levine J. D.*, Journ. Appl. Phys., **23**, 67 (1961).
169. *Goodman A. M., Warfield G.*, Phys. Rev., **120**, 1142 (1960).
170. *Mess C. E. K., James T. H.* (eds.), The Theory of the Photographic Process, New York, 1966.
171. *Seiwatz R., Green M.*, Journ. Appl. Phys., **29**, 1034 (1958).
172. *Levine J. D.*, Surface Sci., **10**, 313 (1968).
173. *Pines D.*, Elementary Excitations in Solids, New York, 1964. (Имеется перевод: Д. Пайнс, Элементарные возбуждения в твердых телах, изд-во «Мир», 1965.)
174. *Haneman D.*, в книге Proc. Intern. Conf. Phys. Semiconductors, Exeter, England, New York, 1962, p. 842.
175. *Gobeli G. W., Allen F. G.*, Phys. Rev., **127**, 141 (1962).
176. *Allen F. G., Gobeli G. W.*, Phys. Rev., **127**, 150 (1962).
177. *Allen F. G., Gobeli G. W.*, в книге Proc. Intern. Conf. Phys. Semiconductors, Exeter, England, New York, 1966, p. 818.

178. *Aspnes D. E., Handler P.*, Surface Sci., **4**, 353 (1966).
179. *Kobayashi A., Sugiyama K., Arata H., Oda Z.*, Journ. Phys. Soc. (Japan), **16**, 2481 (1961).
180. *Schrieffer J. R.*, Phys. Rev., **97**, 641 (1955). (Имеется перевод: в сб. Проблемы физики полупроводников, ИЛ, 1957, стр. 287).
181. *Handler P., Eisenhour S.*, Surface Sci., **2**, 64 (1964).
182. *Rose A.*, Concepts in Photoconductivity and Allied Problems, New York, 1963.
183. *Pollard E. R., Hartke J. H.*, Bull. Am. Phys. Soc., **14**, 115 (1969).
184. *Van Laar J., Scheer J. J.*, Philips. Res. Rept., **17**, 101 (1962).
185. *Van Laar J., Scheer J. J.*, в книге Proc. Intern. Conf. Phys. Semiconductors, Exeter, England, New York, 1962, p. 827.
186. *Van Laar J., Scheer J. J.*, Surface Sci., **3**, 189 (1965).
187. *Handler P.*, Appl. Phys. Lett., **3**, 96 (1963).
188. *Henzler M.*, Phys. Stat. Sol., **19**, 833 (1967).
189. *Frankl D. R.*, Surface Sci., **4**, 201 (1966).
190. *Frankl D. R.*, Surface Sci., **6**, 334 (1967).
191. *Нестеренко Б. А., Снитко О. В.*, ФТТ, **6**, 2913 (1964).
192. *Нестеренко Б. А., Снитко О. В.*, Surface Sci., **5**, 380 (1966).
193. *Нестеренко Б. А., Снитко О. В., Разумник В. Т.*, Surface Sci., **9**, 407 (1968).
194. *Margoninski Y.*, Phys. Rev., **132**, 1910 (1963).
195. *Shepherd W. B.*, Journ. Vac. Sci. Techn., **6**, 908 (1969).
196. *Arthur J. R., Jr.*, Journ. Appl. Phys., **36**, 3221 (1965).
197. *Gobeli G. W., Allen F. G.*, в книге Semiconductors and Semimetals, vol. 2, New York, 1966.
198. *Fischer T. E.*, Phys. Rev., **142**, 519 (1966).
199. *Fischer T. E.*, Helv. Phys. Acta, **41**, 827 (1968).
200. *Kane E. O.*, Phys. Rev., **127**, 131 (1962); **147**, 335 (1966).
201. *Kindig N. B., Spicer W. E.*, Phys. Rev., **138**, A561 (1965).
202. *Brust D.*, Phys. Rev., **139**, A489 (1965).
203. *Allen F. G., Gobeli G. W.*, Journ. Appl. Phys., **35**, 597 (1964).
204. *Gobeli G. W., Allen F. G.*, Phys. Rev., **137**, 245 (1965).
205. *Scheer J. J., Van Laar J.*, Phys. Lett., **3**, 246 (1963).
206. *Fischer T. E.*, Surface Sci., **13**, 30 (1969).
207. *Губанов А. И.*, Квантовая электронная теория аморфных полупроводников, изд-во АН СССР, 1963.
208. *Somoggia G., Nucciotti A., Chiarotti G.*, Phys. Rev., **144**, 749 (1966).
209. *Harrick N. J.*, Internal Reflection Spectroscopy, New York, 1967. (Имеется перевод: Н. Дж. Харрик, Спектроскопия внутреннего отражения, изд-во «Мир», 1970.)
210. *Chiarotti G., Del Signore G., Nannarone S.*, Phys. Rev. Lett., **21**, 1170 (1968).
211. *Somorjai G. A., Lester J. E.*, Journ. Chem. Phys., **43**, 1450 (1965).
212. *Moesta H.*, Surface Sci., **2**, 267 (1964).
213. *Williams R., Willis A.*, Journ. Appl. Phys., **39**, 3731 (1968).
214. *Gray P. V., Brown D. M.*, Appl. Phys. Lett., **8**, 31 (1966).
215. *Mark P.*, Journ. Phys. Chem. Sol., **26**, 959 (1965).
216. *Haneman D.*, Phys. Rev., **170**, 705 (1968).
217. *Chung M. F., Haneman D.*, Journ. Appl. Phys., **37**, 1879 (1966).

218. Haneman D., Chung M. F., Taloni A., Phys. Rev., 170, 719 (1968).
219. Chung M. F., Bull. Am. Phys. Soc., 14, 787 (1969).
220. Kobayashi A., Oda Z., Kawaji S., Arata H., Sugiyama K., Journ. Phys. Chem. Sol., 14, 37 (1960).
221. Stätz H., de Mars G. A., Davis L., Jr., Adams A., Jr., Phys. Rev., 101, 1272 (1956).
222. Palmer D. R., Morrison S. R., Dauenbaugh C. E., Phys. Rev., 129, 608 (1963).
223. Heiland G., Kuntzmann P., Surface Sci., 13, 72 (1969).
224. Proix F., Haudler P., Surface Sci., 5, 81 (1966).
225. Henisch H. K., Noble W. P., Jr., Surface Sci., 4, 486 (1966).
226. Farnsworth H. E., Schlier R. E., Burger G., Burger R. M., Journ. Appl. Phys., 26, 252 (1955).
227. Schlier R. E., Farnsworth H. E., в книге Semiconductors Surface Physics, Philadelphia, 1957, p. 3.
228. Weber R. E., Peria W. T., Journ. Appl. Phys., 38, 4355 (1967).
229. Mark P., Levine J. D. (eds.), Modern Methods of Surface Analysis, Amsterdam, 1970.
230. Campbell B. D., Farnsworth H. E., Surface Sci., 10, 197 (1968).
231. Campbell B. D., Hauge C. A., Farnsworth H. E., в книге The Structure and Chemistry of Solid Surfaces, New York, 1969, p. 33.
232. Proc. of Symposium on the Structure of Surfaces, Surface Sci., 8, 1 (1967).
233. McRae E. G., Surface Sci., 8, 14 (1967).
234. Lander J. J., Morrison J., Journ. Appl. Phys., 34, 1403 (1963).
235. Farnsworth H. E., Schlier R. E., Dillon J. A., Jr., Journ. Phys. Chem. Sol., 8, 116 (1959).
236. Nosker R., Mark P., Levine J. D., Surface Sci., 19, 291 (1970).
237. Seiwatz R., Surface Sci., 2, 473 (1964).
238. Hansen N. R., Haneman D., Surface Sci., 2, 566 (1964).
239. Palmberg P. W., Peria W. T., Surface Sci., 6, 57 (1967).
240. Haneman D., Roots W. D., Grant J. T. P., Journ. Appl. Phys., 38, 2203 (1967).
241. Taloni A., Haneman D., Surface Sci., 10, 215 (1968).
242. Murkland I., Anderson S., Surface Sci., 5, 197 (1966).
243. Feinstein L. R., Shoemaker D. P., Surface Sci., 3, 294 (1965).
244. Зырянов Г. К., Хоанг А. Т., Вестник Ленинградского университета, Физ. хим., № 10 (1969), стр. 69.
245. Benson G. C., Freeman P. I., Dempsey E., Journ. Chem. Phys., 39, 320 (1963).
246. Tokutaka H., Prutton M., Surface Sci., 11, 216 (1968).
247. Pearson G. L., Read W. T., Jr., Morin F. J., Phys. Rev., 93, 666 (1954).
248. Read W. T., Jr., Phil. Mag., 45, 775 (1954); 46, 111 (1955).
249. Shockley W., Phys. Rev., 91, 228 (1953).
250. Igras E., в книге Proc. Intern. Conf. Phys. Semiconductors, Exeter, England, New York, 1962, p. 832.
251. Farnsworth H. E., Marsh J. B., Toots J., в книге Proc. Intern. Conf. Phys. Semiconductors, Exeter, England, New York, 1962, p. 836.

252. *Henzler M., Heiland G.*, Solid State Commun., **4**, 499 (1966).
253. *Fischer T. E.*, Surface Sci., **10**, 399 (1968).
254. *Vitrikhovskii N. I., Kurik M. V.*, Surface Sci., **3**, 415 (1965).
255. *Thomas M. T., Shimaoka G., Dillon F. A., Jr.*, Surface Sci., **6**, 261 (1967).
256. *Kawaji S., Takishima Y.*, Surface Sci., **1**, 119 (1964).
257. *Ducros P.*, Surface Sci., **10**, 295 (1968).
258. *Tomášek M.*, Surface Sci., **4**, 471 (1966).
259. *Palmer D. R., Morrison S. R., Dauenbaugh C. E.*, Phys. Rev. Lett., **6**, 170 (1961).
260. *Handler P.*, Phys. Rev., **126**, 971 (1962).
261. *Law J. T.*, Journ. Appl. Phys., **32**, 848 (1961).
262. *Heiland G., Lamatsch H.*, Surface Sci., **2**, 18 (1964).
263. *Dillon J. A., Jr., Farnsworth H. E.*, Journ. Appl. Phys., **29**, 1195 (1958).
264. *Dillon J. A., Jr., Oman R. M.*, в книге Proc. Intern. Conf. Semiconductors Phys., Prague, 1961.
265. *Allen F. G.*, Journ. Phys. Chem. Sol., **8**, 119 (1959).
266. *Koutecký J.*, Surface Sci., **1**, 281 (1964).
267. *Pugh D.*, Phys. Rev. Lett., **12**, 390 (1964).
268. *Bartoš I.*, Surface Sci., **15**, 94 (1964).
269. *Jones R. O.*, Phys. Rev. Lett., **20**, 992 (1968).
270. *Gobeli G. W., Allen F. G.*, Surface Sci., **2**, 402 (1964).
271. *Henzler M.*, Journ. Appl. Phys., **40**, 3756 (1969).
272. *Palmer D. R., Morrison S. R., Dauenbaugh C. E.*, Journ. Phys. Chem. Sol., **14**, 27 (1960).
273. *Gobeli G. W., Allen F. G., Kane E. O.*, Phys. Rev. Lett., **12**, 94 (1964).
274. Proc. Intern. Conf. Phys. Semiconductors, Paris, 1964, p. 937.
275. *Banbury P. C., Davies E. A., Green G. W.*, в книге Proc. Intern. Conf. Phys. Semiconductors, Exeter, England, New York, 1962, p. 813.
276. *Barnes G. A., Banbury P. C.*, Proc. Phys. Soc., **71**, 1020 (1958).
277. *Forman R.*, Phys. Rev., **117**, 698 (1960).
278. *Handler P., Portnoy W. M.*, Phys. Rev., **116**, 516 (1959).
279. *Garrett C. G. B., Brattain W. H.*, Phys. Rev., **99**, 376 (1955).
280. *Smoluchowski R.*, Phys. Rev., **60**, 661 (1941).
281. *Wang S., Wallis G.*, Journ. Appl. Phys., **30**, 285 (1959).
282. *Henzler M.*, Bull. Am. Phys. Soc., **13**, 1375 (1968).
283. *Müller K.*, в книге The Structure and Chemistry of Solid Surfaces, New York, 1969, p. 35.
284. *Wolsky S. P., Zdanuk E. J., Shooten D.*, Surface Sci., **1**, 110 (1964).
285. *Aven M., Prener J. S.* (eds.), Physics and Chemistry of II—VI Compounds, Amsterdam, 1964.
286. *Sebenne C., Balkanski M.*, Surface Sci., **5**, 410 (1966).
287. *Fischer T. E.*, Phys. Rev., **139**, 1128 (1965).
288. *Fischer T. E., Allen F. G., Gobeli G. W.*, Phys. Rev., **163**, 703 (1967).
289. *Levine J. D.*, Bull. Am. Phys. Soc., **14**, 787 (1969).
290. *Levine J. D.*, Journ. Vac. Sci. Techn., **6**, 549 (1969).

291. *Levine J. D.*, в книге 2nd Combined Symp. Modern Aspects Vacuum and Thin Film Sci., Bell. Labs., 1969.
292. *Levine J. D.*, в книге Gordon Conf. Surface Sci., New Hampshire, 1969.
293. *Koutecký J., Tomášek M.*, Surface Sci., **3**, 333 (1965).
294. *Long D.*, Energy Bands in Semiconductors, New York, 1968, p. 107.
295. *Fischer T. E.*, Surface Sci., **10**, 274 (1968).
296. *Mead C. A.*, Solid State Electron., **9**, 1023 (1966).
297. *Ruppel W.*, в книге II—VI Semiconducting Compounds, New York, 1967, p. 1260.
298. *Kurtin S., Mead C. A.*, Journ. Phys. Chem. Sol., **29**, 1865 (1968).
299. *Turner M. J., Rhoderick E. H.*, Solid State Electron., **11**, 291 (1968).
300. *Oman R. M.*, Journ. Appl. Phys., **36**, 2091 (1965).
301. *Oman R. M., Priolo M. J.*, Journ. Appl. Phys., **37**, 524 (1966).
302. *Many A., Katzir A.*, Surface Sci., **6**, 279 (1967).
303. *Many A., Shappir J., Shaked U.*, Surface Sci., **14**, 156 (1969).
304. *Shappir J., Many A.*, Surface Sci., **14**, 169 (1969).
305. *Lee R. N., Farnsworth H. E.*, Surface Sci., **3**, 461 (1965).
306. *Freund T., Morrison S. R.*, Surface Sci., **9**, 119 (1968).
307. *Revesz A. G., Zaininger K.*, RCA Rev., **29**, 22 (1968).
308. *Heine V.*, Phys. Rev., **138A**, 1689 (1965).
309. *Memming R., Schwandt G.*, Surface Sci., **5**, 97 (1966).
310. *Pagnia H.*, Surface Sci., **10**, 239 (1968).
311. *Davis J. L.*, Surface Sci., **2**, 33 (1964).
312. *Mark P.*, Journ. Phys. Chem. Sol., **25**, 911 (1964); **26**, 1767 (1965).
313. *Sawamoto K., Toyoda H.*, Rev. Elect. Commun. Lab., **14**, 177 (1966).
314. *Kawaji S., Gatos H. C.*, Surface Sci., **1**, 407 (1964).
315. *Flinn I.*, Surface Sci., **10**, 32 (1968).
316. *Kendall E. J. M.*, Surface Sci., **7**, 486 (1967).
- 317*. *Лифшиц И. М., Пекар С. И.*, УФН, **56**, 531 (1955).
- 318*. *Лаврентьев М. А., Шабат Б. В.*, Методы теории функций комплексного переменного, Физматгиз, 1958.
- 319*. *Степанов В. Е., Чалдышев В. А.*, ФТТ, **12**, 1671 (1970).
- 320*. *Subramanian R., Bhagwa K. V.*, Phys. St. Sol., **B48**, 399 (1971); Journ. Phys. C (Sol. St. Phys.), **5**, 798 (1972).
- 321*. *Garcia-Moliner F., Rubio J.*, Journ. Phys. C (Sol. St. Phys.), **2**, 1789 (1969).
- 322*. *Inglesfield J. E.*, Journ. Phys. C (Sol. St. Phys.), **4**, L104 (1971); Journ. Phys. F (Met. Phys.), **2**, 63 (1972).
- 323*. *Watts C. M. K.*, Journ. Phys. F (Met. Phys.), **3**, 272 (1971); Journ. Phys. C (Sol. St. Phys.), **3**, 1483 (1970); **2**, 66 (1966); **1**, 1237 (1968).
- 324*. *Scherer M., Phariseau P.*, Surface Sci., **18**, 298 (1969).
- 325*. *Лифшиц И. М.*, УФН, **83**, 617 (1964).
- 326*. *Fortsmann F.*, Zs. Phys., **235**, 69 (1970).
- 327*. *Fortsmann F., Heine V.*, Phys. Rev. Lett., **24**, 1419 (1970).
- 328*. *Boudreaux D. S.*, Surface Sci., **28**, 344 (1971).

- 329**. *Gadzik J. W.*, Journ. Vac. Sci. Techn., **9**, 591 (1972).
330*. *Ball G., Morgan D. J.*, Phys. St. Sol., **B50**, 199 (1972).
331*. *Allan G., Lengart P.*, Surface Sci., **30**, 641 (1972).
332*. *Miller D. F., Heron D. L., Haneman D.*, Journ. Vac. Sci. Techn., **9**, 906 (1972).
333*. *Yndurain F., Elices M.*, Surface Sci., **29**, 540 (1972).
334*. *Hoffstein V.*, Surface Sci., **32**, 149, 263 (1972).
335*. *Langreth D. C.*, Phys. Rev., **B5**, 2842 (1972).
336*. *Jones R. O.*, Journ. Phys. C (Sol. St. Phys.), **5**, 1615 (1972).
337*. *Aldredge G. P., Kleinmann L.*, Phys. Rev. Lett., **28**, 1264 (1972).
338*. *Viljoen P. E., Jazzar M. S., Fischer T. E.*, Surface Sci., **32**, 506 (1972).
339*. *Büttner H.*, Surface Sci., **32**, 687 (1972).
340*. *Bertoni C. M., Calandva C.*, Phys. Stat. Sol., **B50**, 527 (1972).
341**. *Stęślicka M., Davison S. G., Srinivasan U.*, Journ. Phys. Chem. Solids, **32**, 1917 (1971).
342**. *Srinivasan U., Davison S. G.*, Physica, в печати (1973).
343**. *Fairbairn W. M., Glasser M. L., Stęślicka M.*, Surface Sci., в печати (1973).
344**. *Davison S. G.*, Phys. Lett., **43A**, 67 (1973).
345**. *Stęślicka M.*, будет опубликовано.
346**. *Phariseau P.*, Simon Stevin, **38**, 3 (1964).
347**. *Davison S. G., Tan K. P.*, Surface Sci., **27**, 297 (1971).
348**. *Kolar M., Bartos I.*, Czech. Journ. Phys., **B23**, 179 (1973).
349**. *Davison S. G.*, будет опубликовано.
350**. *Levine J. D.*, Surface Sci., **34**, 90 (1973).
351**. *Goldstein B.*, Surface Sci., **35**, 227 (1973).
352**. *Martinelli R. U.*, Appl. Phys. Lett., **17**, 313 (1970).
353**. *Levine J. D., Willis A., Bottoms W. R., Mark P.*, Surface Sci., **29**, 144 (1972).
354**. *Baidyaroy S., Mark P.*, Surface Sci., **30**, 53 (1972).
355**. *Baidyaroy S., Bottoms W. R., Mark P.*, Surface Sci., **29**, 165 (1972).
356**. *Lidow D., Bottoms W. R.*, Journ. Electrochem. Soc., **119**, 242c (1972).
357**. *Lang B., Toyner R. W., Somorjai G. A.*, Surface Sci., **30**, 440, 454, (1972).

ОГЛАВЛЕНИЕ

Предисловие редактора перевода	5
Предисловие авторов к русскому изданию	6
Глава 1. Введение	7
Глава 2. Краткий исторический очерк	9
Глава 3. Методы кристаллических орбиталей	12
§ 1. Прямой подход	12
§ 2. Методы резольвенты	29
Глава 4. Методы кристаллического потенциала	48
§ 3. Модель Кронига — Пенни	48
§ 4. Задача Матье	79
Глава 5. Расчеты трехмерных случаев	111
§ 5. Непосредственное обобщение одномерной задачи	112
§ 6. Аналитическое продолжение	113
§ 7. Структура приповерхностной области	114
§ 8. Плотность состояний	118
Глава 6. Полуклассические методы	119
§ 9. Приближение ненасыщенных связей	119
§ 10. Метод потенциала Маделунга	123
Глава 7. Сравнение теории и эксперимента	131
§ 11. Основные явления	133
§ 12. Распределение поверхностных состояний	140
§ 13. Экспериментальные методы	151
§ 14. Приготовление поверхности	160
§ 15. Метод дифракции электронов низких энергий (ДЭНЭ)	162
§ 16. Линейные дефекты	165
§ 17. Результаты и их интерпретация	166
Глава 8. Заключительные замечания	218
Литература	222

С. Дэвисон, Дж. Левин

ПОВЕРХНОСТНЫЕ (ТАММОВСКИЕ) СОСТОЯНИЯ

Редактор В. Ф. Трифонов

Художник Ф. Л. Лейн. Художественный редактор Е. Урусов

Технический редактор Ф. Х. Третьякова. Корректор М. М. Колотилина

Сдано в набор 30/1 1973 г. Подписано к печати 13/VI 1973 г.

Бумага тип. № 3 84×108¹/₃₂ = 3,63 бум. л. 12,18 усл. печ. л. Уч.-изд. л. 11,24

Изд. № 2/6739. Цена 84 коп. Заказ № 520

ИЗДАТЕЛЬСТВО «МИР», Москва, 1-й Рижский пер., 2

Ордена Трудового Красного Знамени Ленинградская типография № 2 имени Евгении Соколовой Союзполиграфпрома при Государственном комитете Совета Министров СССР по делам издательств, полиграфии и книжной торговли. г. Ленинград, Л-52, Измайловский проспект, 29