А.И.Ансельм

ВВЕДЕНИЕ В ТЕОРИЮ ПОЛУПРОВОДНИКОВ

Настоящая книга предназначается в первую очередь для лиц, занятых экспериментальными исследованиями в области физики полупроводников. Она, вероятно, окажется полезной и для студентов физических специальностей.

Основное внимание в книге уделено вопросам колебаний кристаллической решетки, законам движения электрона в идеальном и возмущенном периодических полях, кинетическому уравнению и явлениям переноса (прохождению тока).

Для чтения книги требуется знакомство с математикой, квантовой механикой и статистической физикой в объеме программ физического факультета университета или физико-математического факультета политехнического института. При этом не обязательно детальное знакомство с этими курсами, но предполагается, что читатель способен разобраться в соответствующих параграфах учебных книг, если на них делается ссылка.

Особенностью книги является то, что в ней на основе этих простейших сведений все формулы выводятся и, как я надеюсь, достаточно подробно для того, чтобы сделать ее доступной указанному выше кругу лиц.

ОГЛАВЛЕНИЕ

Из предисловия к первому изданию 6
Предисловие ко второму изданию 7
Глава I. Геометрия кристаллических решеток и дифракция рентгеновских лучей 9
§ 1. Простые и сложные кристаллические решетки 9
§ 2. Примеры конкретных кристаллических структур 15
§ 3. Прямая и обратная решетки кристалла 20
§ 4. Формулы Лауэ и Вульфа—Брэгга для дифракции рентгеновских лучей в кристалле. Атомный и структурный факторы рассеяния 24
Глава II. Элементы теории групп и симметрия кристаллов 30
§ 1. Введение 30
§ 2. Элементы абстрактной теории групп 33
§ 3. Точечные группы 39
. § 4. Группа трансляций. Сингонии (кристаллические системы) и решетки Браве 48
§ 5. Кристаллические классы. Пространственные группы 55
§ 6. Неприводимые представления групп и теория характеров 64
§ 7. Квантовая механика и теория групп 79
§ 8. Применение теории групп к исследованию расщепления уровней энергии примесного атома в кристалле и к классификации нормальных колебаний многоатомной молекулы 86
§ 9. Применение теории групп к трансляционной симметрии кристалла 96

§ 10. Правила отбора 106
Глава III. Колебания атомов кристаллической решетки 110
§ 1. Природа сил взаимодействия атомов в кристалле 110
§ 2. Колебания и волны в простой одномерной (линейной) решетке 119
§ 3. Колебания и волны в сложной одномерной (линейной) решетке 125
§ 4. Нормальные координаты для простом одномерной решетки 130
§ 5. Колебания атомов трехмерной сложной кристаллической решетки 133
§ 6. Нормальные координаты колебаний кристаллической решетки 145
§ 7. Колебания простой кубической решетки 151
§ 8. Применение теории групп к исследованию нормальных колебаний кристаллической решетки 157
§ 9. Колебания и волны в кристаллах в приближении изотропного континуума 166
§ 10. Квантование колебаний кристаллической решетки. Фононы 173
§ 11. Теория теплоемкости кристаллической решетки 178
§ 12. Уравнение состояния твердого тела 186
§ 13. Тепловое расширение и теплопроводность твердого тела 191
Глава IV. Электроны в идеальном кристалле 196
§ 1. Общая постановка задачи. Адиабатическое приближение 196
§ 2. Метод Хартри—Фока 199
§ 3. Электрон в периодическом ноле 206
§ 4. Понятие о положительных дырках почти заполненной валентной зоны 217
§ 5. Приближение почти свободных (слабо связанных) электронов 221
§ 6. Зоны Бриллюэна 225
§ 7. Приближение сильно связанных электронов 231
§ 8. Структура энергетических зон и симметрия волновых функций в простой кубической решетке и в кристалле сурмянистого индия 246
§ 9. Группы волнового вектора для решетки типа германия 253
§ 10. Спин-орбитальное взаимодействие и двойные группы 258
§ 11. Двойные группы в кристаллах InSb и Ge 266
§ 12. Спин-орбитал.ьное расщепление в кристаллах InSb и Ge 272
§ 13. Исследование спектра электронов (дырок) вблизи минимума (максимума) энергии в зоне Бриллгоэна (**kp**-метод) 276
§ 14. Симметрия, связанная с обращением времени 291
§ 15. Структура энергетических зон некоторых полупроводников 299
Глава V. Локализованные состояния электрона в кристалле 304
§ 1. Функции Ванье. Движение электрона в поле примеси 304
§ 2. Локализованные состояния электрона в неидеальной решетке 311
§ 3. Экситоны 318
§ 4. Поляроны 325
Глава VI. Электрические, тепловые и магнитные свойства твердых тел 336
§ 1. Металлы, диэлектрики н полупроводники 336

§ 2. Статистическое равновесие свободных электронов в полупроводниках и металлах 338
§ 3. Теплоемкость свободных электронов в металлах и полупроводниках 349
§ 4. Магнитные свойства вещества. Парамагнетизм газов и электронов проводимости в металлах и полупроводниках 352
§ 5. Диамагнетизм атомов и электронов проводимости. Магнитные свойства полупроводников 361
§ 6. Циклотронный (диамагнитный) резонанс 372
§ 7. Контакт полупроводника с металлом. Выпрямление 380
§ 8. Свойства p—n-переходов 387
§ 9. Генерация и рекомбинация носителей тока. Квазиуровни Ферми 394
Глава VII. Оптика полупроводников 398
§ 1. Дисперсионные соотношения Крамерса—Кронига 398
§ 2. Межзонное поглощение света, связанное с прямыми переходами 403
§ 3. Межзонные непрямые переходы 417
§ 4. Поглощение света в полупроводниках свободными носителями 426
§ 5. Поляритоны 428
§ 6. Эффект вращения Фарадея 432
§ 7. Теория межзонного поглощения света в квантующем магнитном поле 436
§ 8. Поглощение света в полупроводниках в однородном электрическом поле (эффект Франца—Келдыша) 445
Глава VIII. Кинетическое уравнение и время релаксации для электронов проводимости в кристаллах 454
§ 1. Явления переноса и кинетическое уравнение Больцмана 454
§ 2. Кинетическое уравнение для электронов в кристалле 463
§ 3. Рассеяние электронов на акустических колебаниях решетки 467
§ 4. Время релаксации электронов проводимости в атомном полупроводнике и металле 471
§ 5. Теория деформационного потенциала в кубических кристаллах с простой зонной структурой 476
§ 6. Рассеяние электронов проводимости в ионных кристаллах на колебаниях решетки 481
§ 7. Рассеяние электронов проводимости на заряженных и нейтральных атомах примесей 488
Глава IX. Кинетические процессы (явления переноса) в полупроводниках 494
§ 1. Введение 494
§ 2, Определение неравновесной функции для электронов проводимости в случае сферически-симметричной зоны 497
§ 3. Электропроводность невырожденных полупроводников с простой зонной структурой 502
§ 4. Термоэлектрические явления в невырожденных полупроводниках с простой зонной структурой 506

§ 5. Гальваномагнитные явления в невырожденных полупроводниках с простой зонной структурой 513
§ 6. Термомагнитные явления в невырожденных полупроводниках с простой зонной структурой 520
§ 7. Явления переноса в полупроводниках с простой зоной при произвольном вырождении 527
§ 8. Явления переноса в полупроводниках типа германия и кремния 533
§ 9. Явления переноса в полупроводниках со сферической непараболической зоной 553
§ 10. Эффект «фононного увлечения» в полупроводниках 557
§ 11. Квантовая теория гальвано- и термомагнитных явлений в полупроводниках , 565
Приложения 574

ИЗ ПРЕДИСЛОВИЯ К ПЕРВОМУ ИЗДАНИЮ

Настоящая книга предназначается в первую очередь для лиц, занятых экспериментальными исследованиями в области физики полупроводников. Она, вероятно, окажется полезной и для студентов физических специальностей.

Основное внимание в книге уделено вопросам колебаний кристаллической решетки, законам движения электрона в идеальном и возмущенном периодических полях, кинетическому уравнению и явлениям переноса (прохождению тока).

Для чтения книги требуется знакомство с математикой, квантовой механикой и статистической физикой в объеме программ физического факультета университета или физико-математического факультета политехнического института. При этом не обязательно детальное знакомство с этими курсами, но предполагается, что читатель способен разобраться в соответствующих параграфах учебных книг, если на них делается ссылка.

Особенностью книги является то, что в ней на основе этих простейших сведений все формулы выводятся и, как я надеюсь, достаточно подробно для того, чтобы сделать ее доступной указанному выше кругу лиц.

Некоторые математические выводы, более сложные и менее связанные с основным текстом, приведены в конце книги в приложениях.

Естественно, что подробный вывод основных соотношений при ограниченном объеме книги заставил отказаться от изложения ряда важных вопросов и от подробного сравнения теории с экспериментом.

Книга эта мною посвящена светлой памяти Абрама Федоровича Иоффе, по инициативе которого она была написана.

Выражаю искреннюю признательность редактору Г. Е. Пикусу за многочисленные замечания, во многом способствовавшие улучшению книги.

Посвящается памяти
Абрама Федоровича ИОФФЕ

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

В новом издании книга выходит в значительно дополненном и переработанном виде.

Однако общие цели и предназначение книги остались прежними. Как и 1-е издание, 2-е адресовано в основном тому же кругу читателей — студентам физических специальностей и физикам-экспериментаторам. Это и определяет уровень знаний, необходимый для чтения книги.

Для 2-го издания написаны две новые главы: Элементы теории групп и симметрия кристаллов (гл. II) и оптика полупроводников (гл. VII); кроме того, переработаны старые параграфы и добавлены новые: §§ 6—10 гл. III, §§ 8—14 гл. IV, § 1 гл. V, § 9 гл. VI, § 6 гл. VIII, §§ 2—6, 9, 11 гл. IX и др.

При необозримо возросшем объеме современной физики твердого тела выбор дополнительно включенного материала естественно носил субъективный характер.

Я по-прежнему старался излагать материал так, чтобы «читатель получил некоторое удовольствие от чтения книги и от нашего желания не приводить его в длительное замешательство при фразе «легко может быть показано»» (из предисловия П. Т. Ландсберга к книге Solid State Theory./ ed. by P. T. Landsberg,— London, 1969).

Хотя объем книги возрос почти в полтора раза, многие важные разделы теории полупроводников (сильные электрические поля, неупорядоченные полупроводники, оптика примесных центров, полупроводниковая электроника и др.) не смогли быть включены в книгу.

Поскольку книга является учебным пособием, в ней почти полностью отсутствуют ссылки на оригинальные работы. В соответствующих частях книги даны ссылки на рекомендуемые монографии и учебные пособия; частично они были использованы при работе над книгой. Кроме того, использован материал лекций по теории полупроводников, которые я читал в Ленинградском политехническом институте им. М. И. Калинина и Ленинградском государственном университете им. А. А. Жданова.

Я выражаю свою искреннюю признательность Григорию Евгеньевичу Пикусу, тщательно прочитавшему всю рукопись, за многочисленные советы и замечания, значительно способствовавшие улучшению книги. Я благодарен Ю. Н. Образцову и Г. Л. Биру за обсуждение ряда вопросов, способствовавшее улучшению текста, и А. Г. Аронову за дополнение (пункт 3), написанное к § 8 гл. VII.

В книге принята самостоятельная нумерация параграфов в каждой главе. Параграфы разбиты на пункты. Ссылки на параграфы и формулы из другой главы содержат римскую цифру — номер этой главы.

Все замечания и пожелания, которые будут мною с благодарностью приняты, просьба направлять по адресу: 194 021 Ленинград, К-21, Политехническая ул., 26, Физико-технический институт им. А. Ф. Иоффе АН СССР.

А. Ансельм

ГЛАВА I

ГЕОМЕТРИЯ КРИСТАЛЛИЧЕСКИХ РЕШЕТОК И ДИФРАКЦИЯ РЕНТГЕНОВСКИХ ЛУЧЕЙ

§ 1. Простые и сложные кристаллические решетки

1. Большинство твердых полупроводников и твердые металлы обладают кристаллической структурой, т. е. представляют собой совокупности огромного числа атомов, упорядоченно расположенных в пространстве. Под упорядоченным расположением атомов в пространстве мы понимаем свойство *пространственной периодичности*, или *трансляционной симметрии*, которыми обладает кристаллическая решетка. Иначе говоря, мы предполагаем, что существуют три некомпланарных (т. е. не лежащих в одной плоскости) вектора $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$ таких, что при смещении всего кристалла как целого на любой из этих векторов он совмещается сам с собою. При этом мы отвлекаемся, конечно, от существования теплового движения атомов и наличия у кристалла внешней поверхности. Как будет видно ниже, направления векторов $\boldsymbol{a}_i$ ($i=1, 2, 3$) могут быть выбраны в решетке различным образом. Кроме того, очевидно, что смещение кристалла на векторы, кратные $\boldsymbol{a}_i$, тоже приводит к совмещению его с самим собой. В дальнейшем мы под $\boldsymbol{a}_i$ будем понимать наименьшие по длине векторы при их фиксированном направлении. При таком выборе величин $\boldsymbol{a}_i$ они называются *трансляционными*, *масштабными* или *основными векторами*, или *трансляционными периодами* кристаллической решетки. Параллелепипед, построенный на трех векторах $\boldsymbol{a}_i$, называется *элементарной* или *кристаллической ячейкой*. Условимся располагать векторы $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$ в такой же последовательности, как и положительные оси x, y и z в правой координатной системе. Пользуясь обычным определением векторного произведения в правой координатной системе, можно показать [1]), что объем элементарной ячейки

$$\Omega_0 = (\boldsymbol{a}_1[\boldsymbol{a}_2\boldsymbol{a}_3]) = (\boldsymbol{a}_3[\boldsymbol{a}_1\boldsymbol{a}_2]) = (\boldsymbol{a}_2[\boldsymbol{a}_3\boldsymbol{a}_1]). \qquad (1.1)$$

[1]) Смирнов В. И. Курс высшей математики.—21 изд.—М.: Наука, 1974, т. 2, п. 117.

2. Мы начнем изучение геометрии кристаллических решеток с рассмотрения линейной (одномерной) решетки, т. е. совокупности частиц, периодически расположенных вдоль бесконечной прямой линии. Такая решетка может быть получена посредством последовательного смещения вдоль прямой линии атома или группы атомов на равные отрезки a. В случае линейной решетки мы имеем только один трансляционный вектор $|\boldsymbol{a}_1| = a$, и «объем» элементарной ячейки Ω_0 равен длине отрезка a. На рис. I.1 представлены три линейные решетки. Белые и черные кружки изображают атомы различного сорта. Учитывая, что трансляционный период $\boldsymbol{a}$ есть наименьшее расстояние, на которое надо сместить решетку, чтобы она совпала сама с собой, мы видим, что решетка $а$) содержит один атом в элементарной ячейке $\Omega_0 = a$, а решетки $б$) и $в$) — по два атома. Решетка $а$) называется *простой* или *примитивной*, а решетки $б$) и $в$) — *сложными*.

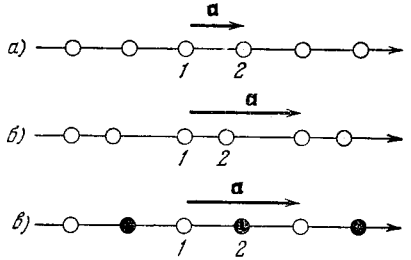


Рис. I. 1.

На рис. I.2, $а$ изображена плоская решетка с атомами, расположенными в вершинах параллелограммов. Она может быть

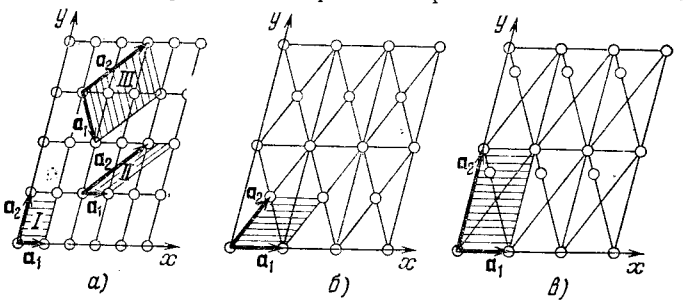


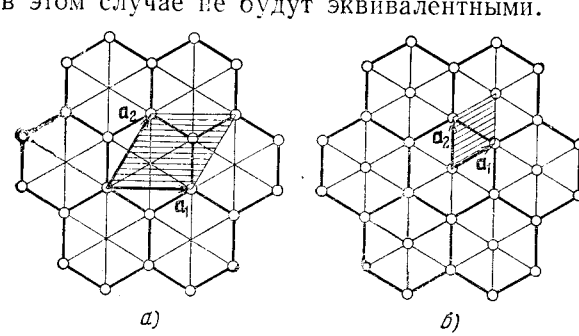
Рис. I. 2.

получена в результате параллельного смещения в плоскости на равные расстояния простой линейной решетки (рис. I.1, $а$). Как показывает рис. I.2, $а$ выбор трансляционных векторов $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$ неоднозначен. Элементарные ячейки *I* и *II* содержат по одному атому и «объем» их $\Omega_0 = |[\boldsymbol{a}_1\boldsymbol{a}_2]|$, равный площади заштрихованных параллелограммов, одинаков. На примере элементарной ячейки *III*, содержащей три атома, мы видим, что и в простой решетке может быть сконструирована элементарная ячейка, содержащая более одного атома. Если основные векторы $\boldsymbol{a}_i$ вы-

браны так, что любая трансляция решетки может быть представлена как $\sum_i n_i \boldsymbol{a}_i$ с целочисленными значениями n_i, то элементарная ячейка, построенная на $\boldsymbol{a}_i$, называется *примитивной*. Элементарные ячейки *I* и *II* на рис. I.2, *а* примитивные, а ячейка *III* не примитивная. В самом деле, в последнем случае смещение вдоль оси x на одну (две) минимальные трансляции равно $n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2$ с $n_1 = n_2 = {}^1/_3 = (= {}^2/_3)$. Если примитивная ячейка содержит один атом, то решетка называется *простой*, если больше одного атома, то — *сложной*; это определение распространяется и на трехмерную решетку.

Таким образом, решетка на рис. I.2, *а* простая.

На рис. I.2, *б* изображена другая простая решетка, которая получается из решетки, приведенной на рис. I.2, *а*, если на пересечении диагоналей параллелограммов поместить атомы того же сорта. Примитивная ячейка может быть теперь выбрана так, как это показано на рис. I.2, *б*. Если мы сдвинем одинаковым образом все атомы, находящиеся на пересечениях диагоналей (см. рис. I.2, *в*), то получим уже сложную решетку с двумя атомами на примитивную ячейку, которую мы можем выбрать так, как это показано на рисунке. Эту сложную решетку мы можем себе представить как две одинаковые простые решетки, вдвинутые одна в другую. Если мы на пересечениях диагоналей параллелограммов (рис. I.2, *б*) поместим атомы другого сорта, то мы получим сложную решетку, так как узлы решетки в этом случае не будут эквивалентными.



а) *б)*

Рис. I.3.

На рис. I.3, *а* изображена весьма симметричная плоская решетка, атомы которой помещены в вершинах шестиугольников, заполняющих плоскость. Нетрудно убедиться в том, что эта решетка сложная, так как примитивная ячейка, изображаемая на рисунке векторами $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$, содержит два атома. Если же дополнительно в центре каждого шестиугольника поместить такой же атом, то мы придем к простой решетке (I.3, *б*).

3. В случае трехмерной кристаллической решетки в ее основе лежит элементарная ячейка в форме параллелепипеда, построенного на основных векторах $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$ (рис. 1.4; через α_{12} обозначен угол между $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$ и т. д.).

Основные векторы решетки могут быть выбраны бесконечным числом способов. Пусть новые («штрихованные») основные векторы

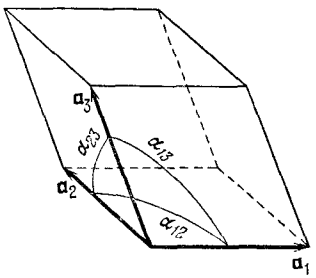


Рис. 1.4.

$$\boldsymbol{a}_i' = \sum_k \beta_{ik} \boldsymbol{a}_k, \tag{1.2}$$

где β_{ik} — целые числа. Для того чтобы вектор решетки

$$\boldsymbol{a} = n_1 \boldsymbol{a}_1 + n_2 \boldsymbol{a}_2 + n_3 \boldsymbol{a}_3, \tag{1.3}$$

где n_1, n_2, n_3 — целые числа, выражался через $\boldsymbol{a}_i'$ формулой, аналогичной (1.3), необходимо и достаточно, чтобы определитель

$$|\beta_{ik}| = \pm 1. \tag{1.4}$$

В самом деле, из (1.2) следует, что

$$\boldsymbol{a}_k = \sum_i \beta_{ki}^{-1} \boldsymbol{a}_i', \tag{1.5}$$

где коэффициенты обратного преобразования [1])

$$\beta_{ki}^{-1} = \frac{A(\beta_{ik})}{|\beta_{ik}|}. \tag{1.6}$$

Здесь $A(\beta_{ik})$ — алгебраическое дополнение (i, k)-го члена определителя $|\beta_{ik}|$.

Подставляя (1.5) и (1.6) в (1.3), получим

$$\boldsymbol{a} = \sum_k n_k \boldsymbol{a}_k = \sum_{k,\ i} n_k \frac{A(\beta_{ik})}{|\beta_{ik}|} \boldsymbol{a}_i'.$$

Для того чтобы обеспечить целочисленность коэффициентов при $\boldsymbol{a}_i'$ для любых n_k, необходимо потребовать выполнение условия (1.4).

4. Мы уже указывали на то, что если примитивная ячейка содержит один атом, то решетка называется простой. Сложные решетки, когда на примитивную ячейку приходится два или более атомов, реализуются как в решетках с одним сортом атомов, так и при наличии нескольких сортов атомов. Рассмотрим несколько основных простых решеток.

На рис. 1.5 изображены три кубические решетки, для которых основные векторы $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$ взаимно перпендикулярны

[1]) Смирнов В. И. Курс высшей математики, — 10 изд. — М.: Наука, 1974, т. 3, ч. 1, гл. II.

($\alpha_{12}=\alpha_{23}=\alpha_{13}=90°$) и имеют одинаковую длину $a_1=a_2=a_3=a$. Про такие решетки говорят, что они принадлежат к *кубической системе*. Решетка на рис. I.5, *а* называется простой кубической. Элементарные ячейки на рис. I.5, *б*, *в* называются *объемноцентрированным кубом* и *гранецентрированным кубом*. В первом случае дополнительный атом помещен в центре куба, во втором случае — в центрах шести боковых граней куба.

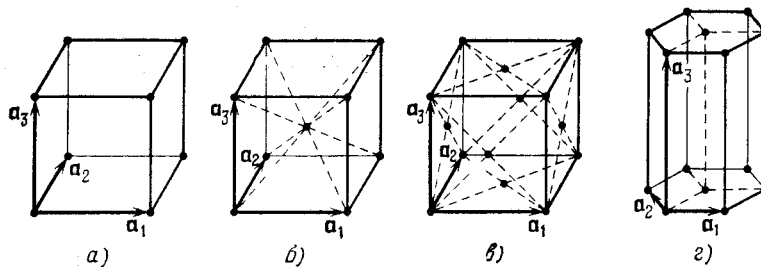


Рис. I.5.

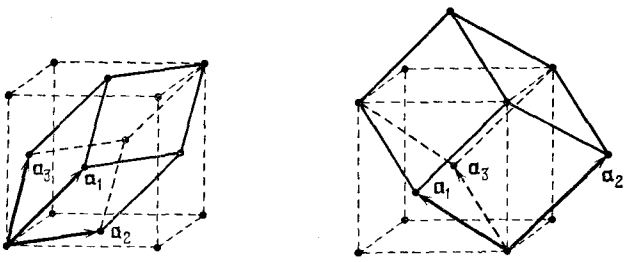
Подсчитаем число атомов в элементарных ячейках кубических решеток. Для этого учтем, что в вершине куба соприкасаются восемь элементарных ячеек (кубов), поэтому на одну ячейку приходится $^1/_8$ атома; атом, расположенный на боковой грани ячейки, вносит в нее $^1/_2$ атома (попутно заметим, что атом, который был бы расположен на ребре ячейки, вносил бы в нее $^1/_4$ атома). Мы видим, что в случае простого куба элементарная ячейка содержит $^1/_8\times 8=1$ атом; в случае объемноцентрированного куба — $^1/_8\times 8+1=2$ атома; в случае гранецентрированного куба — $^1/_8\times 8+^1/_2\times 6=4$ атома. Однако объемноцентрированная и гранецентрированная решетки являются простыми, так как их примитивные ячейки содержат по одному атому. В случае куба с центрированными гранями (когда на всю кубическую ячейку приходится 4 атома) основные векторы примитивной ячейки могут быть направлены из вершины куба к центрам прилегающих граней, как это изображено на рис. I.6, *а*. В этом случае в примитивной ячейке ($\alpha_{12}=\alpha_{23}=\alpha_{13}=60°$) содержится один атом. В объемноцентрированной кубической решетке основные векторы примитивной ячейки можно направить из вершины куба к центрам соседних кубов, как это изображено на рис. I.6, *б*.

На рис. I.5, *г* изображена так называемая *гексагональная ячейка* в виде правильной шестигранной призмы, боковые ребра которой перпендикулярны к основанию правильного шестиугольника. В простой решетке одинаковые атомы расположены в вершинах призмы и в центрах оснований. В гексагональной

решетке можно направить основные векторы примитивной ячейки по трем ребрам шестигранной призмы, сходящимся в одной вершине.

Разделение решеток на простые и сложные представляется существенным при изучении колебаний атомов в кристаллах, так как только сложные решетки обладают оптическими ветвями колебаний.

В общем случае элементарная ячейка, имеющая форму параллелепипеда, не обладает симметрией кристаллической решетки. Так, например, на рис. I.3, *б* повороты плоской



а) *б)*

Рис. I.6.

решетки вокруг любого атома на угол в 60° приводят решетку к самосовпадению, в то время как примитивная ячейка, заштрихованная на рис. I.3, *б*, такой симметрией не обладает. Очевидно, что примитивная ячейка гранецентрированного куба, изображенная на рис. I.6, *а*, не обладает симметрией куба.

Вигнер и Зейтц показали, как выбрать примитивную ячейку так, чтобы она обладала симметрией кристаллической решетки. Возьмем некоторый атом решетки *O* и проведем из него отрезки к ближайшим атомам; построим через середины этих отрезков плоскости, перпендикулярные к ним. Пересечения этих плоскостей определят некоторый минимальный многогранник, содержащий внутри узел *O*; этот многогранник называется *ячейкой Вигнера—Зейтца*. Очевидно, такими ячейками можно плотно заполнить все пространство кристалла [1]).

Если провести эту процедуру для плоской решетки на рис. I.3, *б*, то ячейка Вигнера—Зейтца будет иметь форму правильного шестиугольника, обладающего симметрией гексагональной решетки.

Ячейка Вигнера—Зейтца для простой кубической решетки имеет форму куба. Как будет выглядеть ячейка Вигнера—Зейтца

[1]) Сложная решетка состоит из ряда одинаковых простых решеток, вдвинутых друг в друга; в этом случае указанное построение выполняется для одной из простых подрешеток.

для объемноцентрированной кубической решетки? Выберем в качестве узла O атом в центре куба. Восемь перпендикулярных плоскостей, проведенных через середины отрезков, соединяющих O с атомами в восьми вершинах куба, образуют правильный восьмигранник (*октаэдр*). Шесть перпендикулярных плоскостей, проведенных через середины отрезков, соединяющих O с центральными атомами соседних кубов, отсекут шесть вершин октаэдра образуя четырнадцатигранник, изображенный на рис. IV.5, *а*. Восемь граней его — правильные шестиугольники, а шесть граней — квадраты. Четырнадцатигранная ячейка Вигнера — Зейтца обладает симметрией куба. Аналогично могут быть построены ячейки Вигнера — Зейтца для других кристаллических решеток.

§ 2. Примеры конкретных кристаллических структур

1. Рассмотрим некоторые конкретные кристаллические структуры, иллюстрирующие положения предыдущего параграфа и существенные для дальнейшего изучения теории.

Рентгеноструктурный анализ показывает, что большинство кристаллов чистых металлов принадлежит к кубической или гексагональной системам (рис. I.5, *а*)[1]. Одновалентные щелочные металлы Li, Na, K, Rb и Cs, двухвалентный Ba, переходные металлы, α-, β- и δ-модификации железа и ряд других элементов кристаллизуются в виде объемноцентрированного куба (рис. I.5, *б*). Металлы Cu, Ag, Au, Al, Pb, γ-модификация железа, Ni, Ir, Pt и другие кристаллизуются в форме гранецентрированного куба (рис. I.5, *в*). Элементы Be, Mg, Zn, Cd и другие имеют элементарную ячейку гексагональной структуры (рис. I.5, *г*). Посредством рентгеноструктурного анализа было показано, что в последнем случае мы имеем дело с так называемой плотной гексагональной упаковкой. В этом случае шестигранная призма содержит дополнительно в объеме три атома, как это показано на рис. I.7. При плотной гексагональной упаковке решетка не является простой и содержит два атома в примитивной ячейке.

2. Рассмотрим вопрос о числе ближайших атомов, окружающих данный атом в кристалле и находящихся от него на одном и том же расстоянии d. Это число, называемое *координационным числом*, обозначим буквой z. В любой простой решетке число z одинаково для всех ее узлов. В простой кубической решетке $z=6$ и расстояние d атомов координационной группы до узла равно длине ребра куба a. В объемноцентрированной кубической

[1]) Точное определение кристаллической системы или сингонии будет дано в гл. II, § 4.

решетке $z=8$ и $d=\frac{\sqrt{3}}{2}a$; в гранецентрированной кубической решетке $z=12$ и $d=\frac{\sqrt{2}}{2}a$. В самом деле, для гранецентрированной кубической решетки ближайшими по отношению к вершине куба будут атомы, расположенные в центрах прилегающих граней и отстоящие от вершины на расстоянии $\frac{\sqrt{2}}{2}a$. Очевидно, что в трехмерной решетке в вершине каждого куба сходятся двенадцать таких граней ($z=12$).

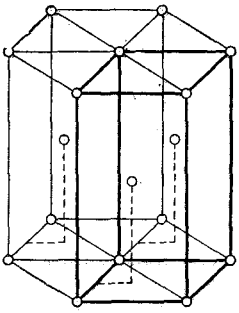
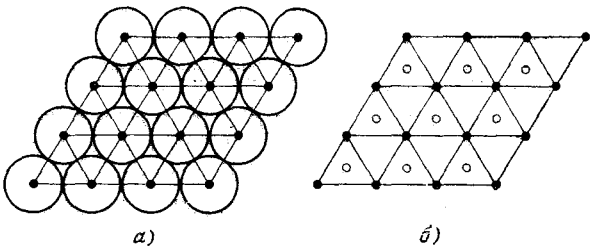


Рис. I.7.



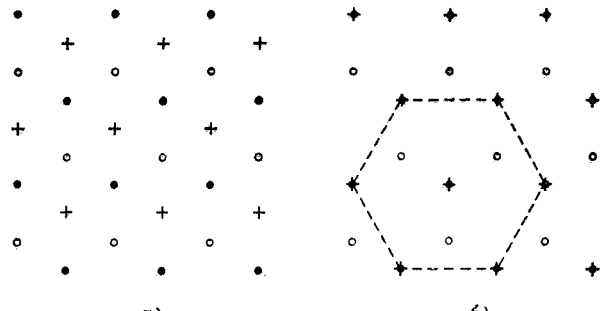
а) *б)*

Рис. I.8.

3. Для рентгеноструктурного анализа имеет значение вопрос о различных плотных упаковках твердых шаров одинакового диаметра. При плотной укладке шаров одинакового диаметра на горизонтальной плоскости центры соприкасающихся шаров располагаются в вершинах равносторонних треугольников со сторонами, равными диаметру шаров (рис. I.8). Второй горизонтальный слой шаров, плотно уложенный на первый, тоже образует сетку таких же равносторонних треугольников. Центры шаров второго слоя расположены по отношению к треугольникам первого слоя так, как это показано кружками на рис. I.8, *б*. Третий горизонтальный слой шаров может быть уложен различно, как это показывает рис. I.9, где ●—атомы первого слоя, ○—атомы второго слоя и +—атомы третьего слоя. На рис. I.9, *б* центры атомов третьего слоя расположены над центрами атомов первого слоя. В этом случае мы имеем дело с плотной гексагональной упаковкой. На рис. I.9, *б* сечение шестигранной призмы гексагональной структуры выделено пунктиром. Можно показать, что упаковке I.9, *а* соответствует структура гранецентрированного куба[1]) (плотная кубическая упаковка).

[1]) Жданов Г. С. Основы рентгеновского структурного анализа.—М.: Гостехиздат, 1940, с. 80.

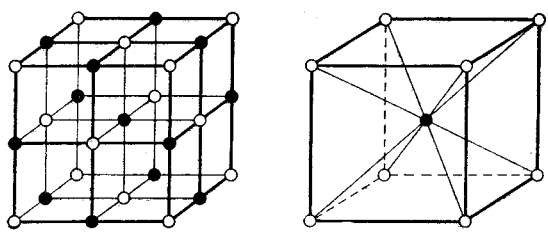
Следует только иметь в виду, что в этом случае ни одна грань куба не будет параллельна горизонтальной плоскости.



а) *б)*

Рис. I. 9.

4. Щелочногалоидные соединения NaCl, LiF, NaI, KCl и т. д., а также бинарные соединения MgO, CaO, MgS, CaSe, BaTe и др. кристаллизуются в форме простой кубической решетки, узлы которой попеременно заняты атомами (точнее ионами) элементов соединения. Такой тип кристаллической решетки носит название структуры каменной соли по имени весьма распространенного соединения NaCl. В этом случае, например, каждый ион Na^+ окружен шестью ионами Cl^- и наоборот (рис. I.10, *а*). Легко видеть, что решетки из Na^+ (или Cl^-)



а) *б)*

Рис. I. 10.

образуют структуру куба с центрированными гранями. Решетки со структурой каменной соли представляют собой сложные решетки с двумя атомами на элементарную ячейку. Трансляционные векторы элементарной ячейки, содержащей два атома, могут быть в случае структуры каменной соли выбраны так же, как в случае простого куба с центрированными гранями (рис. I.6, *а*).

Соединения CsCl, CsBr, CsI имеют структуру объемноцентрированного куба. В этом случае каждый ион Cs^+ окружен восемью отрицательными ионами галоида и аналогично каждый ион галоида — восемью ионами Cs^+ (рис. I.10, *б*).

В рассматриваемых кристаллах число положительных ионов равно числу отрицательных, поэтому как кристалл в целом, так и элементарная кристаллическая ячейка нейтральны.

Легко убедиться в том, что во всех случаях кристаллическая ячейка не только электрически нейтральна, но ее дипольный электрический момент равен нулю. Например, в случае

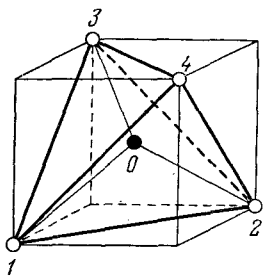


Рис. I.11.

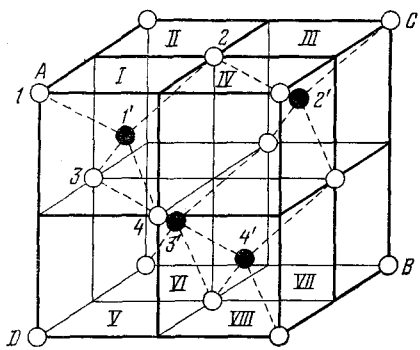


Рис. I.12.

ячейки объемноцентрированного куба (CsCl) электрический момент ячейки, равный нулю, геометрически складывается из восьми дипольных моментов, направленных по четырем объемным диагоналям куба (величина каждого момента равна $\frac{1}{8}e \cdot \frac{a\sqrt{3}}{2}$, где e — заряд одновалентных ионов и a — ребро куба ячейки). Тепловое движение ионов нарушает условия равенства нулю дипольного момента ячейки. Это обстоятельство, как мы увидим ниже, приводит к рассеянию электронов проводимости в ионных кристаллах.

5. За последние годы получили важное техническое применение некоторые вещества (Ge, Si, InSb), имеющие кристаллическую решетку типа алмаза. В решетке этого типа каждый атом, помещенный в центре правильного тетраэдра, окружен четырьмя атомами того же сорта (Ge или Si) или другого (InSb), расположенными в его вершинах. На рис. I.11 изображен узел алмазной решетки *0* с четырьмя окружающими его атомами *1—4*, расположенными в вершинах правильного тетраэдра, вписанного в куб. Легко убедиться в том, что угол между двумя направлениями от узла *0* к окружающим его атомам равен 109°28′.

Решетку алмаза можно также рассматривать как наложение двух кубических гранецентрированных решеток, сдвинутых друг относительно друга в направлении объемной диагонали на $^1/_4$ ее длины. Это хорошо видно на рис. I.12. Пусть вначале мы имели гранецентрированный куб со светлыми кружками ○-атомами. Если мы сдвинем его в направлении объемной диагонали AB на $^1/_4$ ее длины, то ○-атом *1* перейдет на место ●-атома *1′*, ○-атом *2*—на место ●-атома *2′*, ○-атом *3*—на место ●-атома *3′* и ○-атом *4* на место ●-атома *4′*. Так как кубическая гранецентрированная решетка является простой решеткой, то в решетке алмаза можно выделить элементарную ячейку, содержащую два атома.

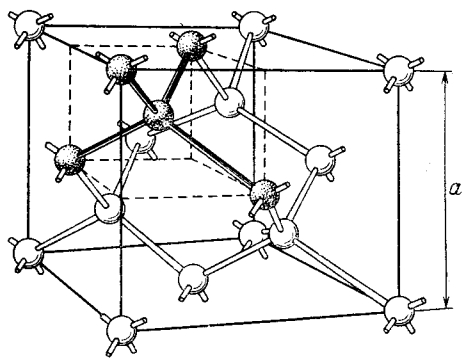


Рис. I.13.

Для германия белые и черные кружки на рис. I.12 изображают атомы одного сорта, но для соединения InSb—разного сорта (например, ○-атомы In и ●-атомы Sb).

Мы можем выделить в алмазной решетке группу из 18 атомов, образующих кубическую ячейку, изображенную на рис. I.13. Расположение атомов в этой группе можно представить себе следующим образом. Разобьем куб с центрированными гранями на восемь одинаковых кубов (*I*—*VIII*) (рис. I.12). Поместим в центры четырех из этих кубов, как это показано на рисунке, атомы (черные кружки). Мы получим решетку типа алмаза, изображенную на рис. I.13. Подсчитаем число атомов в такой кубической ячейке. Из рис. I.12 видно, что 8 атомов расположены в вершинах куба, 6 атомов—на его гранях и 4 атома—в его объеме. Отсюда следует, что число атомов во всей кубической ячейке равно

$$8\cdot\frac{1}{8}+6\cdot\frac{1}{2}+4=8.$$

Не следует думать, что кубическая ячейка (рис. I.12 и I.13), выделенная нами в решетке алмаза, обладает всеми элементами симметрии куба. Так, например, при повороте вокруг вертикальной оси, проходящей через центр куба, на угол 90°, атомы не совмещаются сами с собой. Можно, однако, показать, что по своим макроскопическим свойствам кристалл алмаза обладает кубической симметрией.

§ 3. Прямая и обратная решетки кристалла

1. Главнейшим свойством идеального кристалла, как уже упоминалось выше, является периодическое расположение атомов (точнее атомных ядер) в пространстве. Это означает, что при смещении всего кристалла в целом на вектор

$$\boldsymbol{a} \equiv \boldsymbol{a}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3, \tag{3.1}$$

где $\boldsymbol{a}_i$ — трансляционные векторы решетки, а n_i — целые числа ($i = 1, 2, 3$), кристалл совмещается сам с собой.

Очевидно, что такие величины, как электростатический потенциал или плотность электронов, рассматриваемые в некоторой точке внутри кристалла, являются пространственно- или трехмерно-периодическими функциями. В самом деле, некоторая точка внутри кристалла, определяемая радиусом-вектором $\boldsymbol{r}$, и точка $\boldsymbol{r} + \boldsymbol{a}_n$ физически эквивалентны. Поэтому, например, электростатический потенциал

$$V(\boldsymbol{r}) = V(\boldsymbol{r} + \boldsymbol{a}_n). \tag{3.2}$$

Для разложения периодической функции $V(\boldsymbol{r})$ в ряд Фурье введем координаты ξ_1, ξ_2 и ξ_3 косоугольной системы координат, оси которой направлены по векторам $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$.

В этом случае функция $V(\boldsymbol{r})$ периодична в переменных ξ_1, ξ_2 и ξ_3 с периодами a_1, a_2 и a_3. Разложим периодическую функцию $V(\boldsymbol{r})$ в тройной ряд Фурье [1]), который мы запишем в комплексной форме:

$$V(\boldsymbol{r}) = \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} \sum_{k_3=-\infty}^{+\infty} V_{k_1k_2k_3} e^{2\pi i \left(\frac{k_1\xi_1}{a_1} + \frac{k_2\xi_2}{a_2} + \frac{k_3\xi_3}{a_3} \right)}, \tag{3.3}$$

где k_1, k_2 и k_3 — целые положительные и отрицательные числа, включая нуль.

Переходя в (3.3) от косоугольных координат ξ_i к прямоугольным координатам x_i (см. приложение 1) по формулам

$$\begin{aligned} \xi_1 &= \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3, \\ \xi_2 &= \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3, \\ \xi_3 &= \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3, \end{aligned} \tag{3.4}$$

где α_{ik} — коэффициенты, зависящие от углов между осями косоугольной и прямоугольной координатных систем, получим

$$V(\boldsymbol{r}) = \sum_{b_1} \sum_{b_2} \sum_{b_3} V_{b_1b_2b_3} e^{i(b_1x_1 + b_2x_2 + b_3x_3)}. \tag{3.5}$$

---

[1]) Смирнов В. И. Курс высшей математики, — 21 изд. — М.: Наука, 1974, т. 2, пп. 174, 175; следует иметь в виду, что в книге В. И. Смирнова период обозначается через $2l$, поэтому в показателе экспоненты отсутствует множитель 2.

Здесь b_1, b_2 и b_3 — коэффициенты, зависящие от α_{ik}, k_i, a_i. Суммирование в (3.5) надо вести по всем различным значениям величин b_i, соответствующих всем целочисленным значениям индексов k_i. Рассматривая b_1, b_2 и b_3 как прямоугольные компоненты вектора $\boldsymbol{b}$, запишем (3.5) в виде

$$V(\boldsymbol{r}) = \sum_{\boldsymbol{b}} V_{\boldsymbol{b}} e^{i(\boldsymbol{br})}. \tag{3.6}$$

Проще всего определить $\boldsymbol{b}$ из требования периодичности $V(\boldsymbol{r})$ (3.2):

$$V(\boldsymbol{r}+\boldsymbol{a}_n) = \sum_{\boldsymbol{b}} V_{\boldsymbol{b}} e^{i(\boldsymbol{b},\, \boldsymbol{r}+\boldsymbol{a}_n)} = \sum_{\boldsymbol{b}} V_{\boldsymbol{b}} e^{i(\boldsymbol{br})} e^{i(\boldsymbol{ba}_n)}.$$

Отсюда следует, что $e^{i(\boldsymbol{ba}_n)}$ должно равняться единице, т. е. $(\boldsymbol{ba}_n) = n_1(\boldsymbol{ba}_1) + n_2(\boldsymbol{ba}_2) + n_3(\boldsymbol{ba}_3) = 2\pi \times$ целое число для всех целочисленных значений n_1, n_2 и n_3, что возможно только в том случае, когда

$$\boldsymbol{ba}_1 = 2\pi g_1, \quad \boldsymbol{ba}_2 = 2\pi g_2, \quad \boldsymbol{ba}_3 = 2\pi g_3, \tag{3.7}$$

где g_1, g_2 и g_3 — произвольные целые положительные или отрицательные числа (включая нуль). Всякий вектор определяется тремя своими составляющими (компонентами), поэтому трех независимых уравнений (3.7) достаточно для определения вектора $\boldsymbol{b}$. Можно показать (Приложение 2), что

$$\boldsymbol{b}_g \equiv \boldsymbol{b} = g_1\boldsymbol{b}_1 + g_2\boldsymbol{b}_2 + g_3\boldsymbol{b}_3,$$

где

$$\boldsymbol{b}_1 = \frac{2\pi\,[\boldsymbol{a}_2, \boldsymbol{a}_3]}{\Omega_0}, \quad \boldsymbol{b}_2 = \frac{2\pi\,[\boldsymbol{a}_3, \boldsymbol{a}_1]}{\Omega_0}, \quad \boldsymbol{b}_3 = \frac{2\pi\,[\boldsymbol{a}_1, \boldsymbol{a}_2]}{\Omega_0}, \tag{3.8}$$

а $\Omega_0 = (\boldsymbol{a}_1[\boldsymbol{a}_2, \boldsymbol{a}_3])$ — объем элементарной ячейки.

Легко проверить, что $\boldsymbol{ba}_n = \boldsymbol{b}_g\boldsymbol{a}_n = 2\pi \times$ целое число $= 2\pi(n_1g_1 + n_2g_2 + n_3g_3)$. Непосредственно из определения (3.8) векторов $\boldsymbol{b}_k$ следует, что

$$\boldsymbol{a}_i\boldsymbol{b}_k = 2\pi\delta_{ik} = \begin{cases} 0 & \text{при } i \neq k, \\ 2\pi & \text{при } i = k. \end{cases} \tag{3.9}$$

Наоборот, исходя из уравнений (3.9), можно показать, что векторы $\boldsymbol{b}_k$ определяются выражениями (3.8). Векторы $\boldsymbol{b}_k$ называются *трансляционными*, *масштабными* или *основными векторами обратной решетки*. Векторы $\boldsymbol{a}_n$ и $\boldsymbol{b}_g$ называются векторами *прямой* и *обратной решеток*. Векторы $\boldsymbol{b}_k$ имеют размерность, обратную длине, как это следует из (3.8). Бесконечная периодическая решетка, построенная на трансляционных векторах $\boldsymbol{b}_k$, называется обратной решеткой. Пространство обратной решетки имеет размерность (длина)$^{-3}$. Параллелепипед, построенный на векторах $\boldsymbol{b}_k$, называется элементарной ячейкой обратной

решетки. Можно показать[1]), что его «объем» равен $(\boldsymbol{b}_1[\boldsymbol{b}_2, \boldsymbol{b}_3])=(2\pi)^3/\Omega_0$.

Из условия (3.8) видно, что вектор $\boldsymbol{b}_1$ перпендикулярен векторам $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$, вектор $\boldsymbol{b}_2$ — векторам $\boldsymbol{a}_1$ и $\boldsymbol{a}_3$, вектор $\boldsymbol{b}_3$ — векторам $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$. Если элементарная ячейка прямой решетки имеет форму прямоугольного параллелепипеда, то векторы $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$ параллельны соответственно векторам $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ и $b_i = 2\pi/a_i$. Очевидно, что если прямая решетка простая кубическая, то обратная решетка тоже простая кубическая с $b_1 = b_2 = b_3 = 2\pi/a$.

Представление об обратной решетке возникло непосредственно из задачи разложения в ряд Фурье функции, обладающей периодичностью прямой решетки. В дальнейшем мы увидим, как понятие об обратной решетке плодотворно используется при рассмотрении дифракции рентгеновских лучей в кристаллах, при исследовании колебаний атомов в кристаллах и при квантовомеханическом изучении движения электрона в периодическом поле.

2. Введем важное понятие о миллеровских индексах (hkl), тесно связанное с представлением об обратной решетке.

Представим себе в кристалле плоскость, проходящую через центры атомов. На рис. I.14 изображены четыре такие плоскости, ориентированные различным образом относительно элементарной ячейки простой кубической решетки. Положение

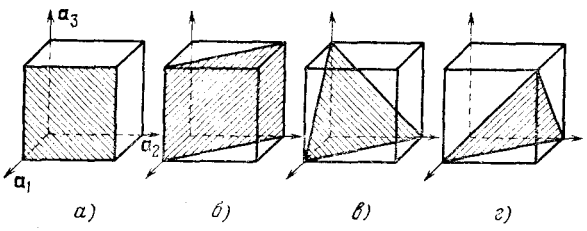


Рис. I. 14.

(ориентацию) плоскости в кристалле, проходящей через атомы, будем характеризовать миллеровскими индексами (hkl), которые определим следующим образом. Пусть целые числа s_1, s_2 и s_3 измеряют в единицах a_1, a_2 и a_3 три отрезка, отсекаемых плоскостью по координатным осям $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$. Составим отношение $\frac{1}{s_1}:\frac{1}{s_2}:\frac{1}{s_3}$ и выразим его через отношение трех наименьших целых чисел. Последние и носят название *миллеровских индексов* (hkl). Таким образом, $h:k:l=\frac{1}{s_1}:\frac{1}{s_2}:\frac{1}{s_3}$. На рис. I.14, *а* изображены оси $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$. Очевидно, что заштрихованная плоскость

[1]) Кочин Н. Е. Векторное исчисление и начало тензорного исчисления.— 9 изд.— М.: Наука, 1965.

на рис. I.14, *а* имеет миллеровские индексы $\frac{1}{s_1}:\frac{1}{s_2}:\frac{1}{s_3}=\frac{1}{1}:\frac{1}{\infty}:\frac{1}{\infty}=$ $=1:0:0$, т. е. $(hkl)=(100)$. Аналогично заштрихованные плоскости на рис. I.14, *б*, *в* и *г* имеют миллеровские индексы (110), (111), (11$\bar{1}$), где $\bar{1}$ означает минус единицу и соответствует тому, что заштрихованная плоскость в случае I.14, *г* отсекает отрезок *а* по отрицательной оси $\boldsymbol{a}_3$. Очевидно, что заданные миллеровские индексы (hkl) определят не одну плоскость, а целое семейство параллельных плоскостей. Легко убедиться, например, в том, что $(\bar{h}\bar{k}\bar{l})=(hkl)$.

Совокупность физически эквивалентных плоскостей, например всех шести граней куба, обозначается символом $\{100\}$. Направление прямой в кристалле, проходящей через центры атомов, указывается символом $[uvw]$, где u, v и w — три наименьших целых числа, отношение которых $u:v:w$ равно отношению длин (в единицах a_1, a_2 и a_3) составляющих вектора по осям $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$, направленного вдоль данной прямой. Так, например, ось x на рис. I.14, совпадающая с вектором $\boldsymbol{a}_1$, имеет символ [100]. Направлению объемной диагонали куба, проходящей через начало координат, соответствует символ [111]. Физически эквивалентные направления в кристалле обозначаются символом $\langle uvw\rangle$. В кубическом кристалле направление $[uvw]$ перпендикулярно плоскости (uvw), что не имеет места в кристаллах более низкой симметрии.

Докажем два важных положения. Построим в пространстве прямой решетки, узлы которой определяются векторами $\boldsymbol{a}_i$, обратную решетку, узлы которой определяются векторами $\boldsymbol{b}_i$, ориентированными относительно $\boldsymbol{a}_i$ согласно формулам (3.8). Конечно, масштаб векторов $\boldsymbol{b}_i$ остается при этом произвольным, так как они имеют размерность $см^{-1}$. Докажем теперь, что вектор обратной решетки $\boldsymbol{b}_g=g_1\boldsymbol{b}_1+g_2\boldsymbol{b}_2+g_3\boldsymbol{b}_3$ перпендикулярен плоскости с миллеровскими индексами (hkl), если $g_1:g_2:g_3=h:k:l$. Очевидно, что концы векторов $\boldsymbol{a}_1/h$, $\boldsymbol{a}_2/k$ и $\boldsymbol{a}_3/l$ лежат на плоскости (hkl), поэтому векторы $\left(\frac{\boldsymbol{a}_1}{h}-\frac{\boldsymbol{a}_2}{k}\right)$ и $\left(\frac{\boldsymbol{a}_1}{h}-\frac{\boldsymbol{a}_3}{l}\right)$ лежат в плоскости (hkl). Возьмем вектор $\boldsymbol{b}_{hkl}=h\boldsymbol{b}_1+k\boldsymbol{b}_2+l\boldsymbol{b}_3$, параллельный вектору $\boldsymbol{b}_g$, и докажем, что $\boldsymbol{b}_{hkl}$ перпендикулярен плоскости (hkl). Для этого достаточно убедиться в том, что скалярные произведения

$$\boldsymbol{b}_{hkl}\left(\frac{\boldsymbol{a}_1}{h}-\frac{\boldsymbol{a}_2}{k}\right)=(h\boldsymbol{b}_1+k\boldsymbol{b}_2+l\boldsymbol{b}_3)\left(\frac{\boldsymbol{a}_1}{h}-\frac{\boldsymbol{a}_2}{k}\right)=0,$$

$$\boldsymbol{b}_{hkl}\left(\frac{\boldsymbol{a}_1}{h}-\frac{\boldsymbol{a}_3}{l}\right)=(h\boldsymbol{b}_1+k\boldsymbol{b}_2+l\boldsymbol{b}_3)\left(\frac{\boldsymbol{a}_1}{h}-\frac{\boldsymbol{a}_3}{l}\right)=0,$$

как это следует из (3.9).

Докажем, во-вторых, что расстояние d_{hkl} между соседними плоскостями семейства (hkl) равно $2\pi/b_{hkl}$. Пусть $\boldsymbol{n}=\boldsymbol{b}_{hkl}/b_{hkl}$ —

единичный вектор нормали к плоскостям (hkl). Тогда

$$d_{hkl}=\left(\frac{\boldsymbol{a}_1}{h}\boldsymbol{n}\right)=\frac{1}{h}\frac{1}{b_{hkl}}\boldsymbol{a}_1(h\boldsymbol{b}_1+k\boldsymbol{b}_2+l\boldsymbol{b}_3)=\frac{2\pi}{b_{hkl}}.$$

Проверим, что расстояние между плоскостями (100) в кубической решетке равно постоянной решетки $\boldsymbol{a}$. Действительно, в этом случае $b_{hkl}=b_{100}=b_1=2\pi/a$ и, следовательно, $d_{100}=2\pi/b_{100}=a$. Так же просто может быть вычислено расстояние между плоскостями (111) в кубической решетке: $b_{111}=\sqrt{b_1^2+b_2^2+b_3^2}=2\pi\sqrt{3}/a$, т. е. $d_{111}=2\pi/b_{111}=a/\sqrt{3}$.

§ 4. Формулы Лауэ и Вульфа — Брэгга для дифракции рентгеновских лучей в кристалле. Атомный и структурный факторы рассеяния

1. Рассмотрим условия дифракции рентгеновских лучей в кристалле по Лауэ. На рис. I.15 изображены два атома O и A, разделенные трансляционным вектором $\boldsymbol{a}_1$. Падающий в направлении $\boldsymbol{n}$ $(n=1)$ пучок рентгеновских лучей рассеивается атомами по всем направлениям. Рассмотрим рассеянный пучок RS в таком направлении, определяемом единичным вектором $\boldsymbol{n}'$ $(n'=1)$, для которого лучи IAR и KOS усиливают друг друга в результате интерференции. Для этого необходимо, чтобы их геометрическая разность хода $BO+OC$ равнялась целому числу волн, т. е. $BO+OC=-(\boldsymbol{a}_1\boldsymbol{n})+(\boldsymbol{a}_1\boldsymbol{n}')=\boldsymbol{a}_1(\boldsymbol{n}'-\boldsymbol{n})=g_1\lambda$, где g_1 — произвольное целое число. В трехмерной решетке условие интерференционного усиления пучка должно одновременно выполняться для всех атомов кристалла, соответствующих трансляциям $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$. Это даст следующие условия интерференции Лауэ:

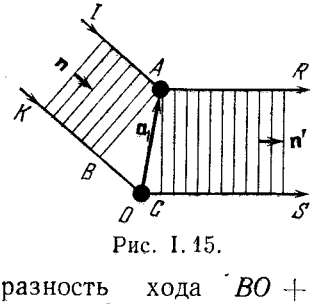


Рис. I.15.

$$\boldsymbol{a}_1(\boldsymbol{n}'-\boldsymbol{n})=g_1\lambda,\quad \boldsymbol{a}_2(\boldsymbol{n}'-\boldsymbol{n})=g_2\lambda,\quad \boldsymbol{a}_3(\boldsymbol{n}'-\boldsymbol{n})=g_3\lambda, \tag{4.1}$$

которые могут быть записаны в более привычном виде:

$$a_1(\cos\alpha_1'-\cos\alpha_1)=g_1\lambda,\quad a_2(\cos\alpha_2'-\cos\alpha_2)=g_2\lambda,$$
$$a_3(\cos\alpha_3'-\cos\alpha_3)=g_3\lambda, \tag{4.2}$$

если ввести углы α_i и α_i' между направлениями падающего и рассеянного лучей и кристаллическими осями $\boldsymbol{a}_i$, т. е. положить $\boldsymbol{a}_i\boldsymbol{n}=a_i\cos\alpha_i$ и $\boldsymbol{a}_i\boldsymbol{n}'=a_i\cos\alpha_i'$.

Из (4.1) следует, что вектор $2\pi\left(\frac{\boldsymbol{n}'}{\lambda}-\frac{\boldsymbol{n}}{\lambda}\right)$ удовлетворяет тем же уравнениям (3.7), что и вектор $\boldsymbol{b}$. Отсюда следует, что

$$2\pi\left(\frac{\boldsymbol{n}'}{\lambda}-\frac{\boldsymbol{n}}{\lambda}\right)=\boldsymbol{b}_g, \tag{4.3}$$

где $\boldsymbol{b}_g = g_1\boldsymbol{b}_1 + g_2\boldsymbol{b}_2 + g_3\boldsymbol{b}_3$ — вектор обратной решетки. Условие (4.3) можно рассматривать как эквивалентное уравнениям Лауэ (4.1) или (4.2).

Введем, по определению, волновой вектор $\boldsymbol{k} = \frac{2\pi}{\lambda}\boldsymbol{n}$. В этом случае условие (4.3) приобретает вид

$$\boldsymbol{k}' - \boldsymbol{k} = \boldsymbol{b}_g. \tag{4.4}$$

Так как $k' = k$, то из (4.4) следует: $\boldsymbol{k}^2 = \boldsymbol{k}'^2 = (\boldsymbol{b}_g + \boldsymbol{k})^2 = \boldsymbol{b}_g^2 + \boldsymbol{k}^2 + 2(\boldsymbol{b}_g\boldsymbol{k})$, откуда

$$^1/_2\boldsymbol{b}_g^2 + (\boldsymbol{b}_g\boldsymbol{k}) = 0, \tag{4.5}$$

что, конечно, также эквивалентно (4.1) и (4.3).

2. Покажем, как могут быть в пространстве обратной решетки геометрически интерпретированы условия Лауэ в форме (4.4). На рис. I.16 изображена для наглядности плоская обратная решетка. Построим из произвольного узла O обратной решетки вектор $-\boldsymbol{k}$ и из начала его R опишем окружность радиуса $k = 2\pi/\lambda$ (сфера Эвальда). Если эта окружность (сфера) пройдет непосредственно вблизи другого узла S обратной решетки, то вектор $\overrightarrow{RS}$ равен $\boldsymbol{k}'$ — волновому вектору рассеянного пучка, усиленного интерференцией. При этом вектор обратной решетки $\overrightarrow{OS}$ равен $\boldsymbol{b}_g$, удовлетворяющему равенству (4.4). Таким образом, конструируя сферу Эвальда в пространстве обратной решетки, мы можем определить графически все направления, по которым будут наблюдаться интерференционные максимумы.

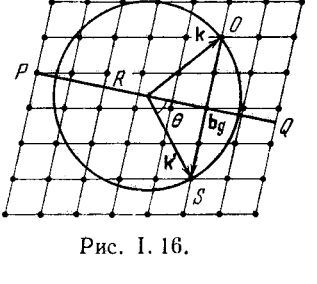


Рис. I. 16.

3. Условия интерференции рентгеновских лучей Лауэ (4.1)— (4.5) могут быть сформулированы в ином виде. Проведем прямую (плоскость) PQ, перпендикулярную вектору обратной решетки $\boldsymbol{b}_g = \overrightarrow{OS}$ (рис. I.16). Как было показано в конце предыдущего параграфа, эта плоскость в пространстве прямой решетки определяет атомную плоскость с миллеровскими индексами (hkl), где $h:k:l = g_1:g_2:g_3$. Обозначая общий множитель чисел g_1, g_2 и g_3 через n, имеем $b_g = nb_{hkl}$. С другой стороны, расстояние между соседними плоскостями (hkl) равно $d = 2\pi/b_{hkl} = 2\pi n/b_g$. Рассмотрим теперь треугольник ORS. Обозначая угол $\sphericalangle QRS$ через θ, имеем $OS = 2RS\sin\theta$ или $b_g = 2k\sin\theta = 2\frac{2\pi}{\lambda}\sin\theta$. Вводя в последнее равенство величину d, получим формулу Вульфа — Брэгга

$$2d\sin\theta = n\lambda. \tag{4.6}$$

Из рис. I.16 видно, что падающий пучок $\boldsymbol{k}$ и рассеянный $\boldsymbol{k}'$ формально можно рассматривать как зеркальное отражение пучка от атомной плоскости PQ с миллеровскими индексами (hkl). Таким образом, условие Вульфа—Брэгга (4.6), эквивалентное условиям Лауэ (4.1), формально можно рассматривать как зеркальное отражение лучей от атомных плоскостей кристалла под углами θ, удовлетворяющими соотношению (4.6).

4. Введем понятие об атомном факторе рассеяния f, учитывающем при рассеянии рентгеновских лучей распределение электрического заряда в атоме. Пусть $\rho(\boldsymbol{r})\,d\tau$—среднее число электронов в объеме атома $d\tau$. Очевидно, что $\int \rho\,d\tau = Z$—полное число электронов в атоме. На рис. I.17 изображен центр атома O и элемент его объема $d\tau$, положение которого в точке A определяется радиусом-вектором $\overrightarrow{OA} = \boldsymbol{r}$. Пусть OS— направление луча, рассеянного одним электроном, который мы мысленно поместим (для простоты расчета) в центре атома O. Амплитуда этого рассеянного излучения в некоторой, достаточно удаленной (по сравнению с размерами атома) точке в заданный момент времени t равна $A(t)$. Считая рассеяние когерентным и учитывая, что его амплитуда пропорциональна числу рассеивающих электронов, получим для амплитуды луча AR в той же точке наблюдения и в тот же момент времени t выражение $dA_\rho(t) = A(t)\,\rho\,d\tau\,e^{i\Phi}$. Здесь $\rho\,d\tau$—число электронов рассеивающего объема $d\tau$ и Φ—разность фаз лучей IAR и KOS. Очевидно, что разность фаз $\Phi = \frac{2\pi}{\lambda} d$, где d—геометрическая разность хода рассматриваемых лучей. Из рис. I.15 видно, что вычисляемая разность хода лучей IAR и KOS, рассеянных атомами O и A, при определении условий интерференции Лауэ, совершенно тождественна с разностью хода лучей, рассеянных центром атома и элементом его объема $d\tau$. Поэтому $d = \boldsymbol{r}(\boldsymbol{n}' - \boldsymbol{n})$, где $\boldsymbol{n}$ и $\boldsymbol{n}'$ — единичные векторы в направлении падающего и рассеянного пучков. Без ограничения общности можно считать, что $\boldsymbol{r}$, $\boldsymbol{n}'$ и $\boldsymbol{n}$ лежат в одной плоскости (чертежа). Построим ось параллельно вектору $\boldsymbol{s} = \boldsymbol{n}' - \boldsymbol{n}$ и плоскость POQ перпендикулярно оси z (см. рис. I.17). Если рассматриваемый нами атом принадлежит кристаллу и направления падающего и рассеянного пучков соответствуют интерференционному максимуму, то POQ есть атомная сетчатая плоскость, от которой происходит зеркальное вульф-брэгговское отражение. В этом случае угол $\sphericalangle POK = \sphericalangle QOS = \theta$ удовлетворяет уравнению Вульфа—Брэгга (4.6). Из рис. I.17

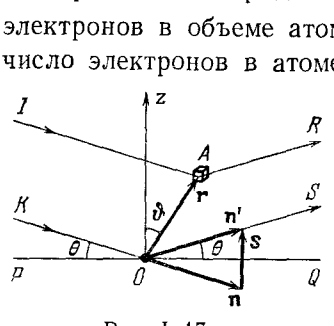


Рис. I.17.

видно, что $s = 2\sin\theta$ ($n = n' = 1$), поэтому $\Phi = \frac{2\pi}{\lambda}(\boldsymbol{rs}) = \frac{2\pi}{\lambda} sr\cos\vartheta = \frac{4\pi}{\lambda}\sin\theta\, r\cos\vartheta$, где ϑ — угол между направлением s (оси z) и радиусом-вектором $\boldsymbol{r}$.

Амплитуда рассеяния всеми электронами атома

$$A_\rho(t) = A(t)\int \rho(\boldsymbol{r})\, e^{i\Phi}\, d\tau.$$

Пусть электроны в атоме распределены сферически симметрично, т. е. $\rho = \rho(r)$: полагая $d\tau = 2\pi r^2 dr \sin\vartheta\, d\vartheta$, получим для атомного фактора рассеяния, по определению,

$$f = \frac{A_\rho(t)}{A(t)} = 2\pi \int_0^\infty \int_0^\pi \rho(r)\, r^2 e^{iar\cos\vartheta} dr \sin\vartheta\, d\vartheta,$$

где $a = \frac{4\pi}{\lambda}\sin\theta$. Интегрируя по ϑ, получим, как показано ниже,

$$f = \int_0^\infty W(r)\frac{\sin(ar)}{ar}\, dr, \qquad (4.7)$$

где $W(r)\, dr = 4\pi r^2 \rho(r)\, dr$ — вероятное (среднее) число электронов в шаровом слое между r и $r + dr$ в атоме.

Для вычисления интеграла $\int_0^\pi e^{iar\cos\vartheta}\sin\vartheta\, d\vartheta$ положим $\cos\vartheta = x$; тогда $dx = -\sin\vartheta\, d\vartheta$ и пределы интеграла равны $+1$ и -1. В результате интеграл равен

$$\int_{-1}^{+1} e^{iarx}\, dx = \frac{1}{iar}\left(e^{iar} - e^{-iar}\right) = \frac{2}{ar}\sin ar.$$

Таким образом, *атомный фактор рассеяния* f[1]) равен отношению амплитуды рассеянного всеми электронами атома излучения по некоторому направлению к амплитуде излучения, рассеянного одним электроном в том же направлении.

Величина $W(r)$, входящая в определение f, может быть вычислена квантовомеханически (метод Харти — Фока) или статистически (метод Томаса — Ферми). На рис. I.18 приведены результаты расчета атомного фактора f для натрия. Для $\theta = 0$ формула (4.7) дает $f = Z$ (число атомных электронов). Экспериментальное изучение распределения интенсивности вблизи интерфе-

[1]) В некоторых случаях (Жданов Г. С. Основы рентгеновского структурного анализа.— М.— Л.: Гостехиздат, 1940, с. 300) величина f называется атомной амплитудой, а атомным фактором называется величина f^2, которой пропорциональна относительная интенсивность рассеянного излучения.

ренционных максимумов рассеяния (при учете температурного фактора рассеяния) подтвердило с большой точностью теоретические расчеты величины f. Эти же исследования при низких температурах подтвердили наличие нулевой энергии колебания атомов в решетке.

5. Мы вывели интерференционные условия Лауэ (4.1) или эквивалентные им условия Вульфа—Брэгга (4.6), рассматривая рассеяние на тождественных атомах простой решетки. Как изменятся результаты, если рассмотреть рассеяние на сложной решетке с базисом, содержащим s атомов, положение которых относительно начала координат ячейки определяется векторами

$$\boldsymbol{r}_n = u_n\boldsymbol{a}_1 + v_n\boldsymbol{a}_2 + w_n\boldsymbol{a}_3 \quad (n = 1, 2, \ldots, s).$$

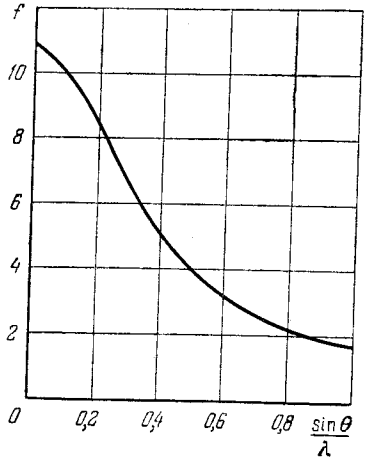


Рис. I. 18.

Так как подрешетка для каждого n-го атома ячейки имеет те же трансляционные векторы $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$, то она дает интерференционные максимумы, определяемые тем же соотношением (4.1), как и для атома с $n = 1$, расположенного в узле (начале координат ячейки). Однако в силу когерентности излучения необходимо учесть фазовые соотношения при рассеянии от разных подрешеток. Введем структурную амплитуду рассеяния $F(hkl)$, равную отношению амплитуды рассеяния в максимуме интерференции (hkl) от всех подрешеток к амплитуде рассеяния в том же направлении от одного электрона.

Если f_n—атомный фактор рассеяния n-го атома, то

$$F(hkl) = \sum_{n=1}^{s} f_n e^{i\Phi_n}, \tag{4.8}$$

где Φ_n—разность фаз для лучей, рассеянных на n-м атоме и атоме, расположенном в начале координат ячейки ($n = 1$). При этом нас интересует интерференционный максимум, определяемый условиями (4.1) с $g_1 = h$, $g_2 = k$ и $g_3 = l$.

Совершенно аналогично вычислению Φ в предыдущем разделе имеем, учитывая (4.1),

$$\Phi_n = \frac{2\pi}{\lambda}\boldsymbol{r}_n(\boldsymbol{n}' - \boldsymbol{n}) = \frac{2\pi}{\lambda}(u_n\boldsymbol{a}_1 + v_n\boldsymbol{a}_2 + w_n\boldsymbol{a}_3)(\boldsymbol{n}' - \boldsymbol{n}) = \\ = 2\pi(hu_n + kv_n + lw_n). \tag{4.9}$$

Из (4.8) и (4.9) следует, что структурный фактор рассеяния, определяющий интенсивность рассеяния, равен

$$|F(hkl)|^2 = \left| \sum_{n=1}^{s} f_n \cos 2\pi (hu_n + kv_n + lw_n) \right|^2 + \\ + \left| \sum_{n=-}^{s} f_n \sin 2\pi (hu_n + kv_n + lw_n) \right|^2. \quad (4.10)$$

Если все атомы кристалла одного сорта, то $f_1 = f_2 = \ldots = f$ и

$$F(hkl) = f \sum_{n=1}^{s} e^{2\pi i (hu_n + kv_n + lw_n)} = fS, \quad (4.11)$$

где

$$S = \sum_{n=1}^{s} e^{2\pi i (hu_n + kv_n + lw_n)}. \quad (4.11a)$$

Рассматривая решетку центрированного куба как сложную, с базисом (0, 0, 0; 1/2, 1/2, 1/2), получим

$$S = 1 + e^{\pi i (h+k+l)}.$$

Если $h+k+l$ — нечетное число, то $e^{i\pi (h+k+l)} = -1$ и $S = 0$, т. е. интенсивность соответствующих максимумов равна нулю.

Таким образом, в объемноцентрированной кубической решетке, состоящей из атомов одного сорта, не будут, например, наблюдаться максимумы интерференции (100), (111), (210), (300) и т. д. Конечно, для решетки объемноцентрированного куба этот же результат можно получить, не пользуясь понятием структурной амплитуды, если выделить элементарную ячейку, соответствующую простой решетке.

ГЛАВА II

ЭЛЕМЕНТЫ ТЕОРИИ ГРУПП И СИММЕТРИЯ КРИСТАЛЛОВ

§ 1. Введение

1. В физику и, в частности, теорию полупроводников все более проникают методы теории групп.

Несмотря на то, что теория групп и ее применение к физике изложены в ряде монографий и учебников[1]), мы сочли необходимым посвятить ей отдельную главу.

Во-первых, большинство книг по теории групп написано слишком сложно и содержит слишком много материала, во всяком случае, для физиков-экспериментаторов. Во-вторых, мы не ставим себе целью обучить читателя применению теории групп, как это делает большинство книг, а хотим только ознакомить его с тем минимумом, который необходим для понимания дальнейших глав книги и современных статей по физике полупроводников.

Теория групп ставит себе целью, исходя только из свойств симметрии физической системы, ответить на ряд вопросов, связанных с ее состоянием. При этом нас будет главным образом интересовать симметрия потенциального поля, действующего на систему (например, симметрия периодического поля, действующего на электроны проводимости в кристалле), а также симметрия, связанная с обращением времени: $t \longrightarrow -t$.

Развивая некоторые общие методы, теория групп позволяет: 1) систематизировать кристаллы по их симметрии, 2) классифицировать нормальные колебания кристалла, 3) классифицировать квантовомеханические состояния электронов проводимости в кристалле и электронов на примесных атомах, 4) определить правила отбора для матричных элементов перехода в системе и др.

[1]) Solid State Theory./Ed. by P. T. Landsberg.—London, 1969, главы IV и V—очень последовательное и понятное изложение; Tinkham M. Group Theory and Quantum Mechanics.—New-York, 1964—одна из лучших книг по применению теории групп в физике (собственно твердому телу посвящена только последняя VIII глава); Бир Г. Л., Пикус Г. Е. Симметрия и деформационные эффекты в полупроводниках,—М., 1972—весьма обширное и глубокое изложение вопроса.

2. Рассмотрим два простых примера, когда следствия, вытекающие из симметрии системы, могут быть получены элементарным путем, без применения аппарата теории групп.

А. Рассмотрим атом гелия, у которого имеются два электрона. Если пренебречь слабым взаимодействием между спинами электронов и их орбитальным движением, то полная волновая функция электронов может быть представлена в виде произведения координатной функции $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ ($\boldsymbol{r}_1$ и $\boldsymbol{r}_2$ — радиусы-векторы электронов) на спиновую функцию $\nu(s_{z1}, s_{z2})$ (s_{z1} и s_{z2} — проекции спинов электронов на ось z). В силу квантовой неразличимости электронов координатная часть волновой функции $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ при перестановке электронов может приобрести только множитель ε, равный по модулю единице. Но в результате дважды произведенной перестановки электронов

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \varepsilon\psi(\boldsymbol{r}_2, \boldsymbol{r}_1) = \varepsilon^2\psi(\boldsymbol{r}_1, \boldsymbol{r}_2),$$

откуда $\varepsilon^2 = 1$ и, следовательно, $\varepsilon = +1$ или $\varepsilon = -1$. Таким образом, мы можем классифицировать состояния электронов в атоме гелия:

$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi(\boldsymbol{r}_2, \boldsymbol{r}_1)$ — *парагелий* с симметричной координатной волновой функцией;

$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = -\psi(\boldsymbol{r}_2, \boldsymbol{r}_1)$ — *ортогелий* с антисимметричной координатной волновой функцией.

Так как полная волновая функция, равная произведению $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)\,\nu(s_{z1}, s_{z2})$, должна быть, в соответствии с принципом Паули, антисимметричной, то спиновая волновая функция $\nu(s_{z1}, s_{z2})$ для парагелия антисимметрична, а для ортогелия — симметрична.

Из симметрии, связанной с перестановкой неразличимых электронов, следует, что парагелию с антисимметричной спиновой функцией соответствует *синглетное* состояние с противоположно направленными спинами и с общим спином, равным нулю; аналогично можно показать, что ортогелию с симметричной спиновой функцией соответствует *триплетное* состояние с параллельными спинами, с суммарным спином, равным $\hbar$.

В самом деле, пусть $\alpha(i)$ и $\beta(i)$ — спиновые функции i-го электрона для спинов, соответственно направленных «вверх» и «вниз»; тогда три симметричные спиновые функции для обоих электронов равны

$$\alpha(1)\alpha(2),\quad \beta(1)\beta(2),\quad \frac{\alpha(1)\beta(2)+\alpha(2)\beta(1)}{\sqrt{2}},$$

где $\sqrt{2}$ введен для нормировки. Можно показать, что полный спин каждого из этих трех состояний равен $\hbar$ (для первых двух это очевидно [1])). С другой стороны, можно составить только одну

[1]) Для третьего см. Блохинцев Д. И. Основы квантовой механики.— 5 изд.— М.: Наука, 1976, § 121.

антисимметричную спиновую функцию, равную $\frac{\alpha(1)\beta(2)-\alpha(2)\beta(1)}{\sqrt{2}}$, с общим спином, равным нулю.

Мы видим, что классификацию электронных состояний атома гелия можно произвести на основании простых соображений симметрии, связанных с неразличимостью электронов.

Б. В качестве второго примера рассмотрим матричный элемент дипольного момента

$$M_x = \int_{-\infty}^{+\infty}\int\int \psi_2^*(\boldsymbol{r})\, x\psi_1(\boldsymbol{r})\, dx\, dy\, dz,$$

возникающий в атоме, если электрическое поле световой волны колеблется вдоль оси x. Здесь $\psi_1(\boldsymbol{r})$ и $\psi_2(\boldsymbol{r})$ — волновые функции валентного электрона в атоме. Вероятность поглощения света в дипольном приближении для перехода $\psi_1 \longrightarrow \psi_2$ пропорциональна квадрату модуля матричного элемента, т. е. $|M_x|^2$. Состояния электрона в сферически симметричном поле атома можно классифицировать по его орбитальному квантовому числу l

$l=0, \ldots, s$ — состояние невырожденное,
$l=1, \ldots, p$ — состояние трижды вырожденное,
$l=2, \ldots, d$ — состояние пятикратно вырожденное

и т. д.; s-состоянию соответствует сферически симметричная волновая функция $\psi_s(r)$ $(r=|\boldsymbol{r}|)$; p-состоянию — три волновые функции, которые могут быть записаны в форме: $x\varphi(r)$, $y\varphi(r)$, $z\varphi(r)$ и т. д. Эта форма волновых функций электрона является непосредственным следствием сферической симметрии действующего на него поля.

Легко видеть, что если оба состояния электрона ψ_1 и ψ_2 являются s-состояниями, то подынтегральное выражение в M_x будет нечетной функцией x, поэтому при интегрировании по x от $-\infty$ до $+\infty$ интеграл будет равен нулю. Если же одно состояние есть s-состояние, а второе — p-состояние, то для волновой функции p-состояния $x\varphi(r)$ подынтегральное выражение — четная функция от всех переменных и, следовательно, $M_x \neq 0$.

Мы приходим к следующим правилам отбора. В дипольном приближении вероятность поглощения света при переходе между s-состояниями равна нулю. Если же переходы совершаются между s- и p-состояниями, то она отлична от нуля.

Теория групп позволяет ответить на вопросы, подобные рассмотренным в пунктах А и Б, и в значительно более сложных случаях.

§ 2. Элементы абстрактной теории групп

1. Группой называется конечное или бесконечное множество различных элементов: A, B, C, ..., P, Q, R ..., удовлетворяющие следующим четырем свойствам.

I. Для каждой пары элементов, взятых в определенной последовательности, определено действие умножения[1]), в результате которого произведению сопоставляется определенный элемент того же множества. Например, произведение элементов B на A нашего множества равно элементу того же множества C, что записывается так: $AB=C$[2]).

Примечание: В общем случае произведение элементов группы непереместительно (некоммутативно), т. е. $AB \neq BA$; если же все элементы группы коммутируют между собою, то группа называется *абелевой* или *коммутативной*.

II. Для произведения элементов группы справедлив сочетательный (ассоциативный) закон, т. е. $(AB)C=A(BC)$.

III. Множество элементов группы должно содержать элемент, называемый *единичным* (мы будем обозначать его E), обладающий свойством $ER=RE=R$, где R — любой элемент группы.

IV. Для каждого элемента группы Q должен среди элементов группы существовать *обратный* элемент (мы будем обозначать его Q^{-1}) такой, что $Q^{-1}Q=QQ^{-1}=E$.

Примечание. Покажем, что $(ABC)^{-1}=C^{-1}B^{-1}A^{-1}$. В самом деле, $(ABC)^{-1}(ABC)=E$: умножая обе части последнего равенства справа на C^{-1}, затем на B^{-1} и затем на A^{-1} и используя сочетательный закон, получим требуемое.

Заметим, что число элементов конечной группы называется ее *порядком*.

2. Примеры групп.

А. Целые положительные и отрицательные числа и нуль образуют бесконечную (абелеву) группу, если под умножением в группе понимать алгебраическое сложение. Единичный элемент в этом случае равен нулю ($E=0$), обратный элемент для элемента (числа) Q равен $-Q$.

По отношению к обычному умножению целые числа группы не образуют, так как обратный элемент $1/Q$ не является целым числом, т. е. не входит в группу.

Б. Множество всех положительных рациональных чисел образует бесконечную абелеву группу относительно обычного арифметического умножения. В этом случае единичный элемент $E=1$, и если $Q=A/B$, то $Q^{-1}=B/A$.

---

[1]) Мы не будем писать слова «умножение», «произведение» в кавычках, хотя соответствующее действие может быть весьма далеким от обычного умножения в алгебре.

[2]) В произведении множители читаются справа налево.

В. Совокупность векторов $\boldsymbol{a}$ образует бесконечную абелеву группу, если умножением в группе является геометрическое сложение векторов: $\boldsymbol{a}+\boldsymbol{a}'=\boldsymbol{a}''$. Единичный элемент $E=\boldsymbol{a}=0$; обратный элемент $(\boldsymbol{a})^{-1}=-\boldsymbol{a}$.

Г. Поучительным примером группы являются вращения равностороннего треугольника *123* (рис. II. 1) с неподвижным центром O, приводящие к совпадению его с самим собой (операции самосовпадения). Высоты треугольника *1a*, *2b* и *3c* пересекаются в центре треугольника O и являются одновременно его биссектрисами и медианами.

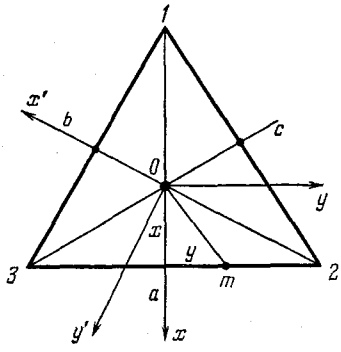


Рис. II. 1.

Введем следующие обозначения для операций самосовпадения треугольника:

1) вращение на 180° вокруг высоты *1a* — A;

2) вращение на 180° вокруг высоты *2b* — B;

3) вращение на 180° вокруг высоты *3c* — C;

4) вращение на угол $120°=2\pi/3$, по часовой стрелке, вокруг оси, перпендикулярной к треугольнику, проходящей через центр O — D;

5) вращение вокруг той же оси и в том же направлении, как и в предыдущем случае, на угол $120°\times 2=240°=4\pi/3$ (это же перемещение можно представить себе как вращение вокруг той же оси против часовой стрелки на отрицательный угол $120°=2\pi/3$) — F;

6) отсутствие перемещения треугольника или вращение его на угол 2π вокруг любых осей — E.

Наряду с рассмотренными операциями вращений имеются также операции симметрии, связанные с отражением в плоскостях, перпендикулярных к плоскости треугольника, проходящих через высоты *1a*, *2в*, *3c*. Мы будем, однако, рассматривать группу 6-го порядка, состоящую из одних вращений A, B, ...

Следует отметить, что операция самосовпадения тела может рассматриваться с двух эквивалентных точек зрения: 1) координатная система (xy) неподвижна, перемещается тело; с этой точки зрения операция D, например, есть вращение *треугольника* на 120° *по часовой стрелке*, так что вершины треугольника переходят друг в друга по схеме: $1\to 2$, $2\to 3$, $3\to 1$; 2) тело неподвижно, но перемещается координатная система: с этой точки зрения операция D есть поворот *координатной системы* на 120° *против часовой стрелки*. Иногда будет удобнее пользоваться

одной точкой зрения, иногда другой. Надо только помнить, что все определяется относительным перемещением тела и координатной системы.

Под произведением AB понимается последовательное применение вначале операции B, затем операции A (аналогично действию операторов на функцию). При этом все оси вращения сохраняют свое положение, не перемещаясь вместе с треугольником. Изобразим произведение AB посредством схемы:

$$AB\left\{\begin{smallmatrix}1\\ \triangle\\ 3\ \ 2\end{smallmatrix}\right\}=A\left\{\begin{smallmatrix}3\\ \triangle\\ 1\ \ 2\end{smallmatrix}\right\}=\begin{smallmatrix}3\\ \triangle\\ 2\ \ 1\end{smallmatrix}=D\left\{\begin{smallmatrix}1\\ \triangle\\ 3\ \ 2\end{smallmatrix}\right\},$$

откуда видно, что это произведение равно D.

Легко проверить, что данное множество из шести элементов удовлетворяет всем I—IV постулатам группы.

Произведение двух элементов группы дает всегда третий. Например, легко проверить, что $AB=D$, $BA=F$, $AA=A^2=E$, $DD=D^2=F$. Последовательное применение операций B и A эквивалентно операции D, а применение этих же операций в обратном порядке эквивалентно операции F. Таким образом, группа неабелева.

Непосредственно из геометрического смысла следует: $A^{-1}=A$, $B^{-1}=B$, $C^{-1}=C$, $D^{-1}=F$, $F^{-1}=D$. В результате мы можем составить следующую таблицу умножения группы (табл. II.1). Эту группу шестого порядка обозначают часто как $\boldsymbol{D_3}$.

Таблица II.1

	E	A	B	C	D	F
E	E	A	B	C	D	F
A	A	E	D	F	B	C
B	B	F	E	D	C	A
C	C	D	F	E	A	B
D	D	C	A	B	F	E
F	F	B	C	A	E	D

← 1-й множитель (правый)

↑ 2-й множитель (левый)

3. Заметим, что в каждой строке и каждом столбце табл. II.1 каждый элемент группы встречается только один раз. Это связано с тем, что если мы имеем группу $\mathfrak{A}=\{A_1\equiv E, A_2, A_3, \ldots, A_h\}$ порядка h и составим произведение из всех элементов группы на один из них, т. е.

$$A_kE,\ A_kA_2,\ A_kA_3,\ \ldots,\ A_kA_h, \tag{2.1}$$

то мы получим все элементы группы, но в ином порядке (*сдвиг по группе*). В самом деле, для этого надо только показать, что все произведения разные, так что не может, например,

$$A_kA_2=A_kA_3;$$

в самом деле, умножим слева обе части этого равенства на A_k^{-1}, тогда $A_k^{-1}A_kA_2=A_k^{-1}A_kA_3$, или $EA_2=EA_3$, или $A_2=A_3$ в противоречии с определением группы.

Если мы будем какой-либо элемент A_k возводить в степень, то будем получать другие элементы группы до тех пор, пока не достигнем степени n, называемой *порядком* элемента A_k, такой, что $A_k^n=E$. В результате получим последовательность элементов

$$A_k,\ A_k^2,\ A_k^3,\ \ldots,\ A_k^n=E. \tag{2.2}$$

При дальнейшем возведении в степень эта последовательность будет повторяться ($A_k^{n+1}=A_k^nA_k=EA_k=A_k$), поэтому она называется *периодом* элемента A_k и обозначается: $\{A_k\}$. Например, для элемента D группы $\boldsymbol{D}_3$ имеем $\{D\}=\{D,\ D^2=F,\ D^3=E\}$.

Очевидно, что период элемента A_k образует абелеву группу порядка n, она называется *циклической*. Вообще, если мы в заданной группе можем выделить некоторое множество элементов, тоже образующих группу, то она называется *подгруппой*. Таким образом, E, D, F—подгруппа группы $\boldsymbol{D}_3$. Тривиальной подгруппой первого порядка (для каждой группы) является единичный элемент E. Так как каждая подгруппа должна содержать единичный элемент, то из группы может быть выделена только одна подгруппа, так как оставшиеся после ее выделения элементы вообще не содержат единичного элемента.

Для конечных групп можно доказать, что порядок подгруппы g есть целый делитель порядка группы h, т. е. h/g—целое число [1]).

Для циклической подгруппы $\{D\}$ группы $\boldsymbol{D}_3$ имеем: $6:3=2$. Отсюда, в частности, следует, что если порядок группы h—простое число, то она имеет только одну (тривиальную) подгруппу E.

---

[1]) Вигнер Е. Теория групп.—М.: ИЛ, 1961, с. 77.

4. Говорят, что элемент B сопряжен элементу A, если

$$B=V^{-1}AV, \tag{2.3}$$

где V—один из элементов группы. Умножая (2.3) на V слева и V^{-1} справа, получим

$$A=VBV^{-1}=(V^{-1})^{-1}BV^{-1}=U^{-1}BU,$$

где $U=V^{-1}$—тоже некоторый элемент группы. Таким образом, если B сопряжен с A, то A сопряжен с B.

Покажем, что если B сопряжен с A и C сопряжен с A, то B сопряжен с C. Имеем

$$B=V^{-1}AV, \quad C=X^{-1}AX,$$

тогда $A=XCX^{-1}$, а

$$B=V^{-1}XCX^{-1}V=(X^{-1}V)^{-1}CX^{-1}V=Z^{-1}CZ,$$

где $Z=X^{-1}V$—тоже некоторый элемент группы. Таким образом, если в (2.3) мы вместо V будем последовательно подставлять все элементы группы, то получим совокупность всех взаимно сопряженных элементов, которая называется *классом*. Определим класс группы $\boldsymbol{D_3}$, содержащий элемент A; элементы, сопряженные с A:

$$E^{-1}AE=A, \quad A^{-1}AA=A, \quad B^{-1}AB=B^{-1}D=C,$$
$$C^{-1}AC=B, \quad D^{-1}AD=C, \quad F^{-1}AF=B,$$

где мы воспользовались табл. II.1. Таким образом, соответствующий класс есть: A, B, C.

Единичный элемент сам по себе образует класс, так как $V^{-1}EV=E$ для любого V. Легко показать, что элементы D и F тоже образуют класс группы D_3. Таким образом, группу D_3 можно разбить на три класса:

$$1)\ E, \quad 2)\ A,\ B,\ C, \quad 3)\ F,\ D. \tag{2.4}$$

Всякая группа может быть разбита на классы, причем один и тот же элемент не может входить в два разных класса.

Каждый элемент абелевой группы образует класс. В самом деле, для абелевой (коммутативной) группы: $V^{-1}AV=V^{-1}VA=A$ для любого V, т. е. каждый элемент сопряжен только с самим собой.

Понятие класса чрезвычайно важно. Оно, конечно, не совпадает с понятием подгруппы. Все классы, кроме E, например классы 2) и 3) в (2.4), не имеют единичного элемента, без которого не может существовать подгруппа.

Подгруппа, состоящая из одного или нескольких целых классов, называется *инвариантной* или *нормальным делителем*. (Вообще говоря, подгруппа не всегда состоит из целого числа классов.)

Если поэтому a—элемент инвариантной подгруппы $\mathcal{H}$ группы G, то все элементы $a'=g^{-1}ag$, где g—любой элемент группы

$G(g \in G)$ принадлежит инвариантной подгруппе (это следует из того, что все элементы класса входят в подгруппу $\mathcal{H}$).

Все элементы одного класса имеют один и тот же порядок n; в самом деле, если $A^n = E$, то $B^n = (V^{-1}AV)^n = V^{-1}AVV^{-1}AV \ldots$ $\ldots V^{-1}AV = V^{-1}A^nV = E$. Легко увидеть, что элементы одного класса обладают некоторым «подобием». Класс 2) в (2.4)—вращения на угол 180°, а класс 3)—вращения на угол 120°. В дальнейшем мы более точно установим сущность «подобия» элементов одного класса.

5. Пусть имеются две, вообще говоря, не абелевы группы

$$\mathfrak{A} = \{A_1 \equiv E, A_2, A_3, \ldots, A_{h_a}\}, \quad \mathfrak{B} = \{B_1 \equiv E, B_2, B_3, \ldots, B_{h_b}\} \tag{2.5}$$

порядков h_a и h_b. Пусть элементы разных групп различны (кроме единичного: $A_1 = B_1 = E$) и коммутируют друг с другом, т. е. $A_kB_l = B_lA_k$ (для всех k и l). Составим h_ah_b произведений A_kB_l и покажем, что это множество—группа.

В первую очередь покажем, что произведение двух элементов множества есть элемент того же множества:

$$(A_{k'}B_{l'})(A_kB_l) = (A_{k'}A_k)(B_{l'}B_l) = A_mB_n,$$

где $A_{k'}A_k = A_m$ и $B_{l'}B_l = B_n$. Произведение A_mB_n есть элемент того же множества.

Покажем, что у множества A_kB_l существует единичный и обратные элементы.

Из (2.5) следует, что единичный элемент множества A_kB_l равен $A_1B_1 = EE = E$. Обратный элемент элемента A_kB_l равен $(A_kB_l)^{-1} = B_l^{-1}A_k^{-1} = A_k^{-1}B_l^{-1}$ и тоже принадлежит рассматриваемому множеству.

Группа из элементов A_kB_l называется *прямым произведением* групп $\mathfrak{A}$ и $\mathfrak{B}$ и обозначается через $\mathfrak{A} \times \mathfrak{B}$. Порядок группы прямого произведения $\mathfrak{A} \times \mathfrak{B}$ равен произведению порядков групп $\mathfrak{A}$ и $\mathfrak{B}$.

Элементы класса группы $\mathfrak{A} \times \mathfrak{B}$, определяемого элементом A_kB_l, равны

$$(A_{k'}B_{l'})^{-1}(A_kB_l)(A_{k'}B_{l'}) = B_{l'}^{-1}A_{k'}^{-1}A_kB_lA_{k'}B_{l'} =$$
$$= (A_{k'}^{-1}A_kA_{k'})(B_{l'}^{-1}B_lB_{l'}) = A_{k''}B_{l''}, \tag{2.6}$$

где мы воспользовались коммутативностью элементов, относящихся к разным группам $\mathfrak{A}$ и $\mathfrak{B}$. В (2.6) $A_{k'}$ и $B_{l'}$ последовательно равны всем элементам групп $\mathfrak{A}$ и $\mathfrak{B}$. Если A_k принадлежит некоторому классу α группы $\mathfrak{A}$, а B_l—некоторому классу β группы $\mathfrak{B}$, то $A_{k''}$ и $B_{l''}$ тоже принадлежат классам α и β. Элемент $A_{k''}B_{l''}$ будет принадлежать классу группы $\mathfrak{A} \times \mathfrak{B}$, происходящему от классов α и β групп $\mathfrak{A}$ и $\mathfrak{B}$. Таким образом, каждой паре классов групп $\mathfrak{A}$ и $\mathfrak{B}$ соответствует класс группы $\mathfrak{A} \times \mathfrak{B}$, следовательно, число классов прямого произведения групп равно произведению числа их классов.

6. Пусть имеются две группы «нештрихованная» и «штрихованная»: $G=\{A, B, C, \ldots, Q, \ldots\}$ и $G'=\{A', B', C', \ldots, Q', \ldots\}$, вообще говоря, с элементами разной природы и разным смыслом. Если между элементами обеих групп можно установить такое взаимно однозначное соответствие: $A \leftrightarrow A'$, $B \leftrightarrow B'$, ..., $Q \leftrightarrow Q'$, ..., чтобы при $BA=Q$ имело место $B'A'=Q'$, т. е. чтобы для обеих групп имелась одинаковая таблица умножения, то группы называются *изоморфными*. Например, элементами группы G могут являться операции движения геометрической фигуры, приводящие ее к совпадению с собой (см. группу равностороннего треугольника $\boldsymbol{D}_3$), а элементами группы G' могут быть матрицы одного ранга, перемножающиеся по той же таблице умножения [1]) (табл. II.1). С точки зрения абстрактной теории групп изоморфные группы тождественны.

Если же нескольким элементам группы G сопоставляется один элемент группы G', например: $A \longrightarrow A'$, $B \longrightarrow A'$, ... $Q \longrightarrow Q'$, при этом из $BA=Q$ следует $A'A'=(A')^2=Q'$, то группа G называется *гомоморфной* группе G'. Гомоморфные группы имеют разный порядок (так как в группе G' все элементы должны быть разными). Тривиальный пример гомоморфизма — это сопоставление всем элементам группы G единицы: $A \longrightarrow 1$, $B \longrightarrow 1, \ldots, Q \longrightarrow 1$ с законом арифметического умножения в группе G'. В этом случае любой группе сопоставляется группа первого порядка ($h=1$).

Если каждой операции симметрии треугольника (E, A, B, ...) на рис. II.1 сопоставить перестановку чисел *1*, *2* и *3*, отмечающих его вершины, то множество перестановок (число которых равно $3!=6$) образует группу, изоморфную $\boldsymbol{D}_3$:

$$E \leftrightarrow (1, 2, 3), \quad A \leftrightarrow (1, 3, 2), \quad B \leftrightarrow (3, 2, 1), \quad \ldots$$

§ 3. Точечные группы [2])

1. Такие перемещения тела конечных размеров (например, молекулы), которые приводят к совпадению его с самим собой (самосовпадению), образуют так называемую точечную группу. Для тела конечных размеров такими перемещениями могут быть повороты на некоторые углы вокруг осей, имеющих общую точку пересечения, и отражения в плоскостях, содержащих эту же

---

[1]) В следующей главе мы подробно рассмотрим этот случай.

[2]) Прекрасное и более полное изложение вопроса можно найти в книге: Ландау Л. Д., Лифшиц Е. М. /Теоретическая физика, т. 3.—Квантовая механика—3 изд.—М.: Наука, 1974, § 91, 93, которой мы близко придерживаемся в этом параграфе.

точку[1]). Ось вращения на угол $2\pi/n$ обозначается через C_n, а на угол $p(2\pi/n)$ через C_n^p; если n кратно p, то $C_n^p = C_{n/p}$. Очевидно, что $C_n^n = E$ = единичному элементу, так как соответствует повороту на угол $n\frac{2\pi}{n} = 2\pi$ или отсутствию поворота, поэтому $C_1 = E$.

Зеркальное отражение в плоскости обозначается символом σ; очевидно, $\sigma^2 = E$ — единичному элементу. Вертикально располагают обычно ось наиболее высокой симметрии (с наибольшим n). Плоскость отражения, проходящую через такую вертикальную ось, обозначают σ_v (vertical), а плоскость отражения, перпендикулярную к такой оси, — через σ_h (horizontal).

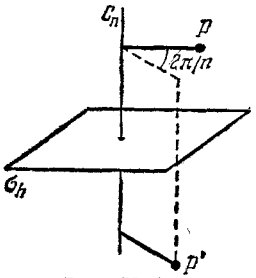


Рис. II.2.

Зеркально-поворотное преобразование (ось) S_n определяется последовательным применением двух операций C_n и σ_h:

$$\sigma_h C_n = C_n \sigma_h = S_n. \tag{3.1}$$

Тело, обладающее зеркально-поворотной симметрией S_n, совпадает само с собой, если повернуть его вокруг оси симметрии на угол $2\pi/n$ и после этого отразить в плоскости, перпендикулярной к оси C_n (или выполнить эти операции в обратном порядке), см. рис. II.2.

Из (3.1) видно, что $S_1 = \sigma_h$. Важный частный случай:

$$\sigma_h C_2 = C_2 \sigma_h = S_2 = J, \tag{3.2}$$

где J — операция *инверсии*. При операции инверсии каждая точка тела P преобразуется в точку P', лежащую на прямой, проходящей через P и неподвижный центр O, так что $\overrightarrow{OP} = -\overrightarrow{OP'}$. При применении оператора инверсии $\hat{J}$ к координатам x, y, z они меняют свой знак, т. е. $\hat{J}\{x, y, z\} = \{-x, -y, -z\}$, поэтому правая координатная система переходит в левую. Из (3.2) следует, что $J\sigma_h = C_2$ и $JC_2 = \sigma_h$; таким образом, три элемента симметрии C_2, σ_h и J взаимно связаны, так что наличие двух из них автоматически приводит к наличию третьего.

2. Укажем ряд геометрических свойств, присущих поворотам вокруг оси и отражениям в плоскости, полезных при изучении точечных групп.

Из кинематики твердого тела хорошо известно, что два последовательных поворота вокруг пересекающихся осей эквива-

[1]) В самом деле, если оси вращения (или плоскости отражения) не имеют общей точки (линии пересечения), то последовательные операции вращений и отражений могут привести к поступательному движению тела, при котором оно совместиться с собой не может.

лентны одному повороту вокруг оси, проходящей через ту же точку [1]). При этом результирующий поворот зависит от порядка, в котором производятся оба поворота (т. е. повороты на конечные углы не коммутируют; поэтому их нельзя изображать векторами).

Последовательное отражение в двух пересекающихся друг с другом плоскостях эквивалентно повороту вокруг оси, совпадающей с прямой пересечения плоскостей, на угол, в два раза бóльший угла между плоскостями φ, т. е.

$$\sigma_v\sigma_v' = C(2\varphi). \qquad (3.3)$$

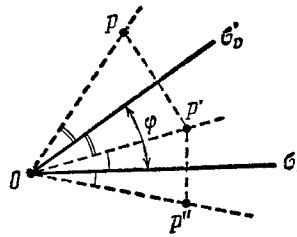


Рис. II.3.

Доказательство этой теоремы непосредственно вытекает из рассмотрения рис. II.3 (заметим, что изменение порядка отражений меняет знак вращения).

Умножая (3.3) слева на σ_v и учитывая, что $\sigma_v^2 = E$, получим

$$\sigma_v' = \sigma_v C(2\varphi), \qquad (3.3а)$$

т. е. произведение поворота и отражения в плоскости, проходящей через ось поворота, эквивалентно отражению в другой плоскости, проходящей через ту же ось и образующую с первой плоскостью угол, равный половине угла поворота. Отсюда вытекает, что если через ось C_n проходит плоскость σ_v, то автоматически возникают еще $n-1$ плоскостей отражения, проходящих через ту же ось, так что углы между ними равны π/n. Аналогично можно показать [2]), что наличие оси C_2, перпендикулярной к оси C_n, автоматически приводит к появлению еще $n-1$ осей C_2, перпендикулярных к C_n, так что углы между ними равны π/n.

Хотя результат двух последовательных преобразований зависит, вообще говоря, от их порядка, в следующих случаях операции коммутативны (переместительны): 1) два поворота вокруг одной и той же оси; 2) два поворота на угол π вокруг взаимно перпендикулярных осей; 3) два отражения во взаимно перпендикулярных плоскостях; 4) поворот и отражение в плоскости, перпендикулярной к оси поворота (см. (3.1)); 5) любой поворот (или отражение) и инверсия в точке, лежащей на оси вращения (или плоскости отражения).

В § 2 п. 4 мы уже упоминали на примере группы $\boldsymbol{D_3}$ о «подобии» элементов, принадлежащих к одному классу. Для

---

[1]) Зоммерфельд А. Механика.—М.: ИЛ, 1947, § 22.

[2]) Ландау Л. Д., Лифшиц Е. М. /Теоретическая физика, т. 3.—Квантовая механика—3 изд.—М.: Наука, 1974, с. 419.

обобщения этого докажем теорему: два поворота на одинаковый угол вокруг разных осей (или два отражения в разных плоскостях) принадлежат к одному классу точечной группы, если среди элементов последней есть операция, совмещающая разные оси поворота (или разные плоскости отражения).

Доказательство. На рис. II.4 изображены оси Oa и Ob; элемент точечной группы B переводит ось Oa в ось Ob; C_n и C_n' — повороты на угол $2\pi/n$ вокруг осей Oa и Ob. Произведение

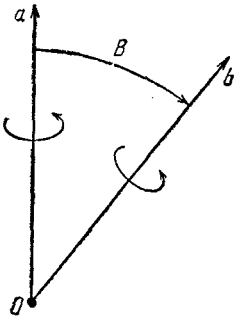


Рис. II.4.

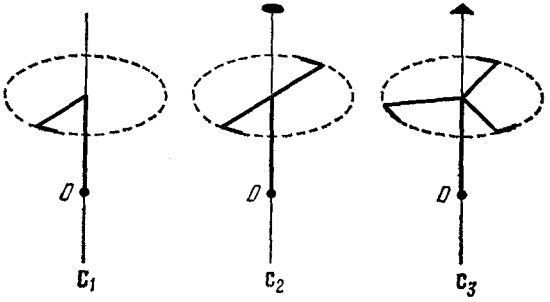


Рис. II.5.

$B^{-1}C_n'B$ имеет следующий геометрический смысл: ось Oa поворачивается до совпадения ее с осью Ob, вокруг последней производится поворот на угол $2\pi/n$, после чего ось Ob поворачивается до совпадения ее с Oa. Очевидно, что результат есть вращение на угол $2\pi/n$ вокруг оси Oa, т. е.

$$C_n = B^{-1}C_n'B,$$

но это и означает, что C_n и C_n' взаимно сопряженные элементы, т. е. принадлежат к одному классу.

Доказательство для отражений в двух разных плоскостях совершенно аналогично. Такие оси и плоскости, которые совмещаются одним из элементов группы, называются *эквивалентными*.

Два поворота вокруг одной и той же оси на одинаковые углы, но в противоположных направлениях, т. е. элементы C_n^k и $(C_n^k)^{-1}$ принадлежат к одному классу, если среди элементов группы имеется ось C_2, перпендикулярная к оси поворота, или плоскость σ_v, проходящая через нее. В самом деле, в этом случае операции C_2 и σ_v меняют направление вращения на противоположное. В этом случае ось вращения C_n называется *двухсторонней*. Заметим, что отражение в плоскости σ_h, перпендикулярной к оси C_n, не изменяет направление вращения и поэтому не делает ось двухсторонней.

Рассмотрим основные точечные группы; при этом мы не будем стремиться ни к полноте, ни к исчерпывающему изложе-

нию вопроса. Символы точечных групп будем изображать жирными латинскими буквами.

I. Группы C_n. Эти группы состоят из одной оси симметрии n-го порядка, т. е. самосовпадение тела имеет место при вращении его вокруг оси на угол $2\pi/n$. Группа циклическая, состоящая из n элементов: C_n, C_n^2, ..., $C_n^n = E$. Каждый элемент является классом. Группа C_1 состоит из одного элемента $C_1 = E$; ей соответствует отсутствие всякой симметрии. На рис. II.5 изображены тела с симметрией C_1, C_2 и C_3 [1]).

II. Группы S_n. Зеркально-поворотная ось S_n при нечетном $n = 2p+1$ сводится к оси симметрии C_{2p+1} и перпендикулярной к ней плоскости симметрии σ_h; в самом деле, $S_{2p+1}^{2p+1} = \sigma_h$, так что в этом случае зеркально-поворотная ось соответствует виду симметрии, рассмотренному в III: группе C_{nh}.

При четном $n = 2p$ группа S_{2p} — циклическая, состоящая из $2p$ элементов: S_{2p}, S_{2p}^2, ..., $S_{2p}^{2p} = E$, каждый из которых является отдельным классом. Группа S_2 состоит из двух элементов: $S_2 = C_2\sigma_h = J$ — инверсии и $S_2^2 = E$ — единичного элемента, она обозначается через C_i. На рис. II.6 изображено тело, имеющее

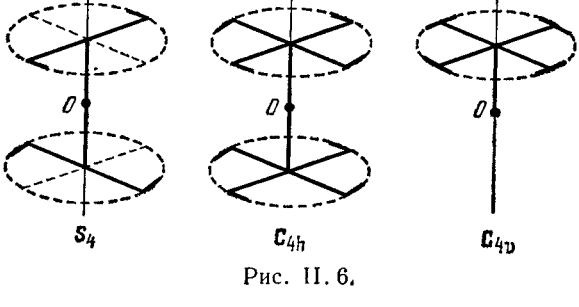


Рис. II.6.

симметрию S_4. В самом деле, при вращении жесткого каркаса (изображенного жирными линиями) на угол $2\pi/4 = 90°$ и отражении в плоскости σ_h, проходящей через центр O, каркас совпадает со своим начальным положением (до вращения).

III. Группы C_{nh}. Эти группы получаются присоединением к оси C_n перпендикулярной к ней плоскости σ_h. Группа содержит $2n$ элементов: n поворотов вокруг оси C_n и n зеркальных поворотов [2]): $C_n^k\sigma_h = S_n^k$, $k = 1, 2, \ldots, n$ (в том числе отражение $C_n^n\sigma_h = \sigma_h$). Все элементы группы коммутативны, т. е. группа абелева; число классов равно числу элементов. Если n четно

[1]) Оси симметрии второго, третьего, четвертого и т. д. порядков обозначаются через ●, ▲, ■ и т. д.

[2]) Группа должна удовлетворять условию: произведение двух элементов группы должно давать тоже элемент группы.

($n=2p$), то группа содержит центр симметрии (инверсию) (т. е. $C_{2p}^{p}\sigma_h=C_2\sigma_h=J$). Простейшая группа C_{1h} содержит два элемента E и σ_h; ее обозначают также через $\boldsymbol{C_s}$. На рис. II.6 изображено тело, обладающее симметрией C_{4h}.

Так как при четном $n=2p$ группа $\boldsymbol{C_{2p,h}}$ содержит инверсию, то группу $\boldsymbol{C_{2p,h}}$ можно представить в виде прямого произведения $\boldsymbol{C_{2p}}\times\boldsymbol{C_i}$, где $\boldsymbol{C_i}=\{E, J\}$.

IV. Группа $\boldsymbol{C_{nv}}$. Эта группа получается, если к оси C_n присоединить проходящую через нее плоскость σ_v. Согласно доказанной выше теореме это автоматически приведет к появлению еще $n-1$ плоскостей, проходящих через ось C_n, так что углы между ними будут равны π/n. Таким образом, группа $\boldsymbol{C_{nv}}$ содержит $2n$ элементов: n вращений вокруг оси C_n и n отражений в плоскостях σ_v. Наличие плоскостей σ_v делает ось C_n двухсторонней.

Распределение по классам различно для четных и нечетных n. Если n нечетно ($n=2p+1$), то последовательные повороты совмещают все плоскости σ_v друг с другом и, следовательно, n отражений принадлежат одному классу. Так как ось C_n двухсторонняя, то каждая пара вращений C_{2p+1}^{k} и $(C_{2p+1}^{k})^{-1}$ ($k=1, 2, \ldots, p$) принадлежит одному классу, общее число которых равно p. Наконец, $C_{2p+1}^{2p+1}=E$ — единичный элемент — тоже отдельный класс. Таким образом, мы имеем всего $p+2$ класса.

Если n четно ($n=2p$), то последовательными поворотами можно совместить только чередующиеся через одну плоскости; в этом случае имеются две системы эквивалентных плоскостей и, следовательно, два класса. Что касается поворотов вокруг оси, то $C_{2p}^{2p}=E$ и $C_{2p}^{p}=C_2$ образуют каждый сам по себе класс, а остальные повороты попарно сопряжены и дают еще $p-1$ класс. Таким образом, полное число классов равно: $2+2+(p-1)=p+3$.

На рис. II.6 изображено тело, обладающее симметрией C_{4v} (две плоскости σ_v проходят через ось симметрии и жирные диаметры; две другие плоскости делят пополам углы, образованные первыми плоскостями).

V. Группы $\boldsymbol{D_n}$. Если к вертикальной оси симметрии C_n добавить перпендикулярную к ней горизонтальную ось C_2, то это автоматически приведет к появлению еще $n-1$ горизонтальных осей C_2, так что углы между ними будут равны π/n (см. выше). Число элементов равно $2n$ (n поворотов вокруг C_n и n вращений вокруг осей C_2). Ось C_n является двухсторонней. Распределение по классам совершенно аналогично группам C_{nv}. На рис. II.7 изображено тело с симметрией $\boldsymbol{D_4}$ (обратите внимание на разницу между $\boldsymbol{D_4}$ и $\boldsymbol{C_{4h}}$); четыре оси C_2 проходят через центр O. Заметим, что группа равностороннего треугольника (§ I.2) есть группа $\boldsymbol{D_3}$.

VI. Группы D_{nh}. Если добавить к системе осей D_n горизонтальную плоскость отражения σ_h, то это автоматически приведет к появлению n вертикальных плоскостей σ_v, проходящих через ось C_n и одну из перпендикулярных к ней осей C_2 (это следует из (3.3а), если положить $\sigma_v = \sigma_h$ и $2\varphi = 180°$). Получающаяся при этом группа D_{nh} содержит $4n$ элементов: $2n$ элементов группы D_n, n отражений σ_v и n зеркально-поворотных преобразований $C_n^k\sigma_h$.

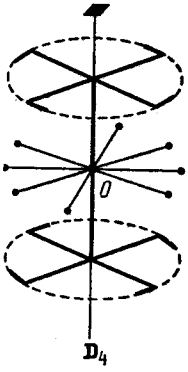


Рис. II.7.

Отражение σ_h коммутирует со всеми остальными элементами группы; поэтому можно написать D_{nh} в виде прямого произведения $D_n \times C_s$ (или $C_{nv} \times C_s$), где C_s есть группа из двух элементов E и σ_h. Отсюда следует, что число классов в группе D_{nh} равно удвоенному числу классов в группе D_n. Половина из них совпадает с классами группы D_n, а другая половина получается умножением их на σ_h. Очевидно, что при четном $n = 2p$, когда группа $D_{2p,h}$ содержит инверсию J, можно написать $D_{2p,h} = D_{2p} \times C_i$.

VII. Группы D_{nd}. Если провести плоскость симметрии через ось C_n группы D_n, так чтобы она делила пополам угол между соседними осями C_2, то это приведет к появлению еще $n-1$ плоскостей, проходящих через ось C_n.

Получающаяся при этом группа D_{nd} содержит $4n$ элементов. К $2n$ элементам группы D_n присоединяются n отражений в вертикальных плоскостях, обозначаемых через σ_d (d — от слова diagonal) и n преобразований вида $C_2\sigma_d$. Можно показать[1]), что $C_2\sigma_d = S_{2n}^{2k+1}$, где $k = 1, 2, \ldots, (n-1)$, т. е. представляют собой зеркально-поворотные преобразования вокруг вертикальной оси, которая оказывается не только осью симметрии n-го порядка, но и зеркально-поворотной осью $2n$-го порядка.

Перейдем теперь к рассмотрению весьма важных точечных групп более высокой симметрии (когда имеется несколько пересекающихся осей n-го порядка), называемых *кубическими*. Называются они так потому, что их элементы симметрии можно выбрать из числа осей и плоскостей симметрии куба.

VIII. Группа осей тетраэдра T. Эта группа состоит из трех осей C_2 и четырех осей C_3 правильного тетраэдра. На рис. II.8, *а, б* показано, как эти оси провести в тетраэдре и как их можно выбрать в кубе (оси C_2 перпендикулярны граням куба, оси C_3 совпадают с его объемными диагоналями);

[1]) Ландау Л. Д., Лифшиц Е. М., Квантовая механика.— М. 1974, стр. 420.

на рис. II.8, *б* в куб вписан тетраэдр с вершиной *a* и основанием *bcd*. Всего имеется 12 элементов: E, $3C_2$, $4C_3$, $4C_3^2 (= 4C_6^{-1})$. Три оси C_2 эквивалентны. Оси C_3 тоже эквивалентны, но не являются

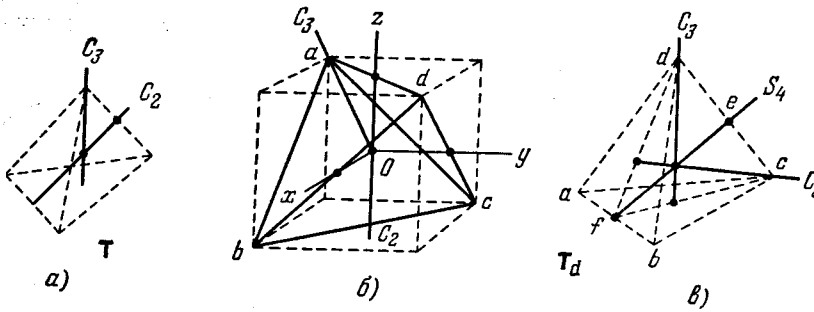


Рис. II.8.

двухсторонними, поэтому элементы группы $\boldsymbol{T}$ распределяются по четырем классам:

$$E;\ 3C_2;\ 4C_3;\ 4C_3^2.$$

Симметрия $\boldsymbol{T}$ не является полной симметрией тетраэдра (см. группу $\boldsymbol{T_d}$). Для того чтобы получить фигуру, обладающую симметрией $\boldsymbol{T}$, достаточно взять произвольное тело, не обладающее симметрией и подвергнутьего всем двенадцати преобразованиям группы $\boldsymbol{T}$; мы получим фигуру, обладающую симметрией $\boldsymbol{T}$.

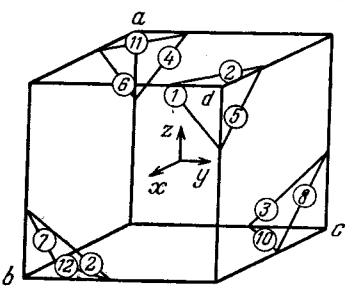


Рис. II.9.

Такая фигура изображена на рис. II.9. Метка *1* последовательно подвергнута всем преобразованиям группы $\boldsymbol{T}$. Повороты C_2 вокруг осей x, y, z преобразуют *1* в *2*, *3*, *4*; повороты C_3 и $C_3^2 = C_3^{-1}$ вокруг осей, проходящих через вершины тетраэдра a, b, c, d, размножают метку по положениям *5*, ..., *12*.

IX. Группа $\boldsymbol{T_d}$. Эта группа содержит все преобразования симметрии тетраэдра. Для того чтобы ее получить, надо к трем осям C_2 и четырем осям C_3 добавить шесть плоскостей отражения, каждая из которых проходит через одну ось C_2 и две оси C_3 (см. рис. II.8, *в*[1])). При этом оси C_2 превращаются в зеркально-поворотные оси S_4. В самом деле, произведение поворота C_2 вокруг оси Ox по часовой стрелке (рис. II.8, *б*) и отражения в плоскости aOd перемещает вершины тетраэдра следующим образом: $a \to c$, $d \to b$, $b \to d \to a$, $c \to a \to d$. С другой

[1]) Вторая плоскость, проходящая через ту же ось C_2 на рис. II.8, *в*, проходит через ребро *afb* и точку *e*.

стороны, зеркальный поворот вокруг оси Oz на угол $\pi/2$ с последующим отражением в плоскости, перпендикулярной к Oz (каждая из этих операций в отдельности не является операцией симметрии тетраэдра!), перемещает вершины тетраэдра так: $b \to a$, $d \to b$, $a \to c$, $c \to d$, что совпадает с результатом первых двух операций. Отсюда следует, что в тетраэдре имеются оси симметрии S_4, направленные вдоль осей C_2. Поскольку плоскости симметрии проходят через оси C_3, последние становятся двухсторонними и, следовательно, элементы C_3 и $C_3^2 = C_3^{-1}$ принадлежат к одному классу. Все плоскости и оси каждого рода эквивалентны. Поэтому 24 элемента группы распределяются по следующим пяти классам[1]):

$$E, \quad 8(C_3, C_3^2) \quad 6\sigma, \quad 6(S_4, S_4^3), \quad 3C_2.$$

X. Группа T_h. Эта группа получается, если к группе T добавить центр симметрии (инверсию), так что $T_h = T \times C_i$. В результате получается вдвое бóльшее число элементов и классов, чем в группе T (т. е. 24 элемента, расположенных по 8 классам).

XI. Группа осей куба O. Это одна из наиболее важных точечных групп: она состоит из всех осей симметрии куба; трех осей C_4, проходящих через центры противоположных граней, четырех осей C_3 — через противоположные вершины и шести

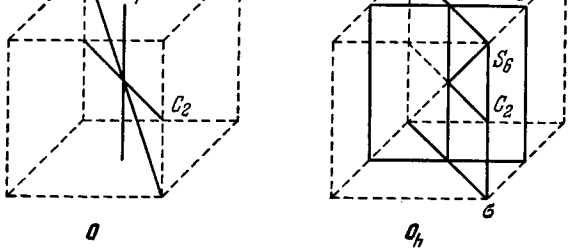


Рис. II.10.

осей C_2 — через середины противоположных ребер (рис. II.10). Все оси одинакового порядка эквивалентны и все оси двухсторонние. Поэтому 24 элемента группы распределяются по следующим пяти классам:

$$E, \quad 8(C_3, C_3^2), \quad 3(C_2 = C_4^2), \quad 6C_2, \quad 6(C_4, C_4^3).$$

Подчеркнем, что группа O не дает полной симметрии куба. Фигура, обладающая симметрией O, может быть получена ана-

[1]) Здесь и дальше численный коэффициент равен полному числу элементов класса, например $8(C_3, C_3^2) = 4C_3 + 4C_3^2$.

логично описанному в группе T, т. е. «размножением» некоторой несимметричной фигуры 24 преобразованиями симметрии группы O. Нетрудно показать, что группа O изоморфна группе тетраэдра T_d.

XII. Группа O_h. Эта группа есть группа полной симметрии куба (октаэдра). Она наряду со всеми осями симметрии куба содержит центр симметрии (инверсию). Так как E и инверсия J коммутируют со всеми элементами группы O, то группа O_h равна прямому произведению групп O и $C_i=\{E, J\}$, т. е. $O_h=O\times C_i$. Можно показать, что группа O_h может быть представлена и как прямое произведение групп T_d и C_i, т. е. $O_h=T_d\times C_i$ [1]). Число элементов группы O_h равно $48\,(24\times 2)$, число классов равно $10\,(5\times 2)$. Все элементы и классы группы O_h могут быть получены умножением элементов и классов группы O (или T_d) на E и J (отсюда следует, что половина элементов и классов группы O_h совпадает с элементами и классами групп O или T_d, а остальные получаются умножением на J). Добавление центра инверсии автоматически приводит к появлению шести плоскостей отражения, проходящих через противоположные ребра (это можно показать, используя (3.2)). Легко также показать, как и в случае группы T_d, что при этом оси C_3 превращаются в зеркально-поворотные оси S_6 (рис. II. 10), а оси 4-го порядка в зеркально-поворотные оси S_4, в результате чего появляются еще три плоскости отражения, перпендикулярные этим осям, т. е. параллельные граням куба.

Такой же симметрией обладают, как мы уже отмечали в гл. I, § 1, ячейки Вигнера—Зейтца для объемноцентрированного куба (четырнадцатигранник, изображенный на рис. IV. 5, *а*) и для гранецентрированного куба (двенадцатигранник или додекаэдр, изображенный на рис. IV. 6, *а*).

§ 4. Группа трансляций. Сингонии (кристаллические системы) и решетки Браве

1. Рассмотрим множество векторов решетки

$$\boldsymbol{a}\equiv\boldsymbol{a}_n=n_1\boldsymbol{a}_1+n_2\boldsymbol{a}_2+n_3\boldsymbol{a}_3, \tag{4.1}$$

соответствующих целочисленным значениям n_1, n_2 и n_3. Множество (4.1) образует группу, если в качестве закона умножения принять геометрическое сложение векторов $\boldsymbol{a}$; единичный элемент группы

$$E=\boldsymbol{a}_n=0. \tag{4.2}$$

Элемент, обратный $\boldsymbol{a}_n$, равен $-\boldsymbol{a}_n$. Эта группа, которую мы будем обозначать через $\mathcal{T}$, называется группой трансляций.

[1]) Более подробно это будет показано в гл. IV, § 8.

Очевидно, что группа $\mathcal{T}$ является абелевой, так что каждый элемент (4.1) является классом группы.

Для идеального бесконечного кристалла трансляционная группа бесконечна. В § 8, п. 1 мы покажем, как бесконечная группа трансляций посредством задания некоторых условий (цикличности) может быть превращена в конечную группу с большим, но конечным числом элементов.

Если в каждый узел $\boldsymbol{n}(n_1, n_2, n_3)$ (4.1) поместить одинаковый атом сферической формы, то мы получим *простую* (или *пустую*) кристаллическую решетку.

Простые решетки могут удовлетворять не только трансляционной симметрии (4.1), но и дополнительно симметрии точечной группы. Например, простая кубическая решетка, удовлетворяющая трансляционной симметрии с тремя взаимно перпендикулярными основными векторами $\boldsymbol{a}_i$ ($i=1, 2, 3$) одинаковой длины, симметрична относительно преобразований точечной группы $\boldsymbol{O}_h$ (начало координат можно поместить в один из узлов решетки или в центр кубической ячейки). Точечную группу симметрии простой решетки мы будем обозначать $\mathcal{F}$. Очевидно, любой элемент R группы $\mathcal{F}$ преобразует любой вектор решетки $\boldsymbol{a}_n$ в другой вектор решетки $\boldsymbol{a}_{n'}$, т. е. $R\boldsymbol{a}_n = \boldsymbol{a}_{n'}$.

Мы покажем, что трансляционная симметрия (4.1) (при произвольных $\boldsymbol{a}_i$) налагает ограничения на группы точечной симметрии $\mathcal{F}$, которым должна удовлетворять простая решетка. Как мы сейчас покажем, для простых решеток имеется только семь точечных групп *(сингоний)*, совместимых с трансляционной симметрией.

В первую очередь отметим, что так как наряду с вектором решетки $\boldsymbol{a}_n$ существует вектор решетки $-\boldsymbol{a}_n$ (для этого достаточно у всех целых чисел n_i поменять знаки на обратные), то простая решетка симметрична (инвариантна) относительно инверсии J. Сложная решетка (см., например, рис. I.2, *в*), вообще говоря, не инвариантна относительно операции инверсии.

Во-вторых, покажем, что с трансляционной симметрией совместимы только оси симметрии 2-го, 3-го, 4-го и 6-го порядков. Оси 5-го, 7-го и более высоких порядков существовать в кристаллах не могут.

Приведем изящное доказательство этого, принадлежащее П. Ниггли (1919).

Проведем в кристалле плоскость, перпендикулярную осям симметрии C_n ($n=2\pi/\alpha$, α—наименьший угол поворота вокруг оси симметрии C_n). Пусть эта плоскость совпадает с плоскостью рис. II.11. A и B—следы двух ближайших друг к другу осей C_n, так что отрезок $\overline{AB}=a_1=$ длина основного вектора решетки (если $\boldsymbol{a}_1$ не лежит в рассматриваемой плоскости, то $\overline{AB}$—проек-

ция a_1 на эту плоскость). Вращая кристалл вокруг осей A и B на угол α в противоположных направлениях, получим равнобокую трапецию $ABA'B'$, в которой $\overline{BA'}=\overline{AB'}=a_1$. Если вращение на угол α—операция симметрии кристалла, то через точки A' и B' тоже проходят оси C_n, поэтому $\overline{B'A'}=pa_1$, где p—целое число (включая нуль).

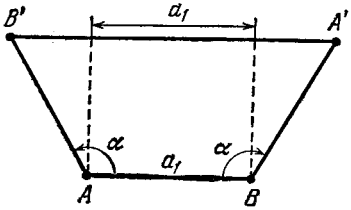


Рис. II.11.

Из рис. II.11 следует, что

$$\overline{B'A'}=a_1+2a_1\sin(\alpha-90^\circ)=pa_1,$$

откуда

$$\cos\alpha=\frac{1-p}{2}. \qquad (4.3)$$

Так как $|\cos\alpha|\leqslant 1$, то возможны только следующие значения $p=3, 2, 1, 0, -1$. Этим значениям p соответствуют следующие углы α:

$$\begin{aligned}
&p=3, && \cos\alpha=-1, && \alpha=180^\circ=2\pi/2,\\
&p=2, && \cos\alpha=-1/2, && \alpha=120^\circ=2\pi/3,\\
&p=1, && \cos\alpha=0, && \alpha=90^\circ=2\pi/4,\\
&p=0, && \cos\alpha=1/2, && \alpha=60^\circ=2\pi/6,\\
&p=-1, && \cos\alpha=1, && \alpha=0^\circ \text{ или } 2\pi.
\end{aligned} \qquad (4.4)$$

Таким образом, в кристалле могут существовать только следующие оси симметрии: C_2, C_3, C_4 и C_6.

Третье ограничение, накладываемое на точечную группу $\mathcal{F}$, заключается в том, что если точечная группа простой решетки содержит ось C_n (где $n>2$), то она удовлетворяет и симметрии C_{nv} [1]).

Если просмотреть все точечные группы (§ 3), то можно установить, что только семь групп, а именно $\boldsymbol{S}_2$, $\boldsymbol{C}_{2h}$, $\boldsymbol{D}_{2h}$, $\boldsymbol{D}_{3d}$, $\boldsymbol{D}_{4h}$, $\boldsymbol{D}_{6h}$ и $\boldsymbol{O}_h$, удовлетворяют всем трем поставленным выше условиям.

Легко видеть, что группы $\boldsymbol{S}_2$, $\boldsymbol{C}_{2h}$ и $\boldsymbol{D}_{2h}$ удовлетворяют поставленным выше требованиям (в этом случае $n=2$, и третье условие не требуется). Далее группы $\boldsymbol{C}_{nv}$ не содержат центра инверсии, а группы $\boldsymbol{C}_{nh}$ ($n>2$) не содержат подгруппу $\boldsymbol{C}_{nv}$. Группы $\boldsymbol{D}_{4h}$ и $\boldsymbol{D}_{6h}$ удовлетворяют всем трем поставленным выше требованиям. Группа $\boldsymbol{D}_{3h}$ не имеет центра инверсии и должна быть заменена на группу $\boldsymbol{D}_{3d}$ [2]). Наконец, из кубических групп поставленным выше условиям удовлетворяет только группа $\boldsymbol{O}_h$.

---

[1]) Доказательство см. у Любарского Г. Я. Теория групп и ее применение в физике.— М., 1957, с. 32.

[2]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика.— М., 1974, с. 420.

Указанные выше семь точечных групп $\mathscr{F}$, образующие семь *кристаллических систем* или *сингоний*, имеют следующие названия и обозначения:

1) триклинная $(\boldsymbol{S}_2)\ldots tr$,
2) моноклинная $(\boldsymbol{C}_{2h})\ldots m$,
3) ромбическая или ортогональная $(\boldsymbol{D}_{2h})\ldots o$,
4) тетрагональная или квадратная $(\boldsymbol{D}_{4h})\ldots t$, (4.5)
5) ромбоэдрическая или тригональная $(\boldsymbol{D}_{3d})\ldots rh$,
6) гексагональная $(\boldsymbol{D}_{6h})\ldots h$,
7) кубическая $(\boldsymbol{O}_h)\ldots c$.

Можно показать, что к одной сингонии могут относиться несколько различных типов простых решеток. Например, к кубической сингонии (c) относятся: простая (Γ_c), объемноцентрированная (Γ_c^v) и гранецентрированная (Γ_c^f) кубические решетки (см. рис. I.5, рис. II.13). Можно показать, что не существует других простых решеток, обладающих точечной симметрией $\boldsymbol{O}_h$.

Триклинная сингония (tr) обладает наинизшей симметрией $\boldsymbol{S}_2=\boldsymbol{C}_i=\{E, J\}$, которой удовлетворяет простая решетка (Γ_t) с основными векторами произвольной длины ($a_1\neq a_2\neq a_3$), произвольно ориентированными по отношению к друг другу.

Французский кристаллограф *А. Браве* (1850 г.) показал, что существует всего 14 типов простых решеток (*решеток Браве*), соответствующих семи сингониям.

Покажем на примере моноклинной сингонии (m), как последовательно определяются возможные типы решеток Браве.

Покажем вначале, что в случае моноклинной сингонии, когда решетка обладает симметрией $\boldsymbol{C}_{2h}$, два основных вектора решетки (например, $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$) лежат в плоскости σ_h, перпендикулярной к оси C_2. Возьмем два вектора решетки $\boldsymbol{a}$ и $\boldsymbol{a}'$, не лежащие в одной плоскости проходящей через ось C_2. Векторы $\boldsymbol{a}+\hat{\sigma}_h\boldsymbol{a}$ и $\boldsymbol{a}'+\hat{\sigma}_h\boldsymbol{a}'$ — тоже векторы решетки ($\hat{\sigma}_h\boldsymbol{a}$ — вектор решетки, получаемый из $\boldsymbol{a}$ отражением его в плоскости σ_h) и, как легко видеть, лежат в плоскости σ_h. Но это значит, что два основных вектора $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$ можно выбрать в плоскости σ_h.

Представим третий основной вектор в виде

$$\boldsymbol{a}_3=\boldsymbol{\alpha}+\boldsymbol{\beta},$$

где $\boldsymbol{\alpha}\parallel C_2$ и $\boldsymbol{\beta}\perp C_2$. Вектор решетки

$$\hat{C}_2\boldsymbol{a}_3-\boldsymbol{a}_3=\hat{C}_2(\boldsymbol{\alpha}+\boldsymbol{\beta})-\boldsymbol{\alpha}-\boldsymbol{\beta}=\boldsymbol{\alpha}-\boldsymbol{\beta}-\boldsymbol{\alpha}-\boldsymbol{\beta}=-2\boldsymbol{\beta}$$

лежит в плоскости σ_h, поэтому

$$2\boldsymbol{\beta}=m_1\boldsymbol{a}_1+m_2\boldsymbol{a}_2$$

(m_1 и m_2 — целые числа или нуль). Таким образом,

$$\boldsymbol{a}_3=\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\alpha}+\frac{m_1}{2}\boldsymbol{a}_1+\frac{m_2}{2}\boldsymbol{a}_2. \qquad (4.6)$$

Легко видеть, что если от $\boldsymbol{a}_3$ отнять вектор $n_1\boldsymbol{a}_1+n_2\boldsymbol{a}_2$ (n_1 и n_2—целые числа), то

$$\boldsymbol{a}_3'=\boldsymbol{a}_3-n_1\boldsymbol{a}_1+n_2\boldsymbol{a}_2$$

является тоже основным вектором решетки (для плоской решетки, в этом можно убедиться, рассматривая рис. I.2, *а*). Положим третий основной вектор равным

$$\boldsymbol{a}_3'=\boldsymbol{a}_3-n_1\boldsymbol{a}_1-n_2\boldsymbol{a}_2=\boldsymbol{\alpha}+\frac{m_1}{2}\boldsymbol{a}_1+\frac{m_2}{2}\boldsymbol{a}_2-n_1\boldsymbol{a}_1-n_2\boldsymbol{a}_2=$$

$$=\boldsymbol{\alpha}+\left(\frac{m_1}{2}-n_1\right)\boldsymbol{a}_1+\left(\frac{m_2}{2}-n_2\right)\boldsymbol{a}_2$$

и выберем коэффициенты при $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$ такими, чтобы они были положительны и не превышали единицу; тогда, в зависимости от того, будут ли m_1 и m_2 четными или нечетными числами, получим следующие четыре случая:

$$\begin{aligned}&1)\ \boldsymbol{a}_3'=\boldsymbol{\alpha};\quad 2)\ \boldsymbol{a}_3'=\boldsymbol{\alpha}+{}^1/_2\boldsymbol{a}_1;\\ &3)\ \boldsymbol{a}_3'=\boldsymbol{\alpha}+{}^1/_2\boldsymbol{a}_1+{}^1/_2\boldsymbol{a}_2;\quad 4)\ \boldsymbol{a}_3'=\boldsymbol{\alpha}+{}^1/_2\boldsymbol{a}_2.\end{aligned}\qquad(4.7)$$

Легко видеть, что 2) отличается от 4) только переобозначением векторов $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$, а 4) переходит в 3), если заменить основной

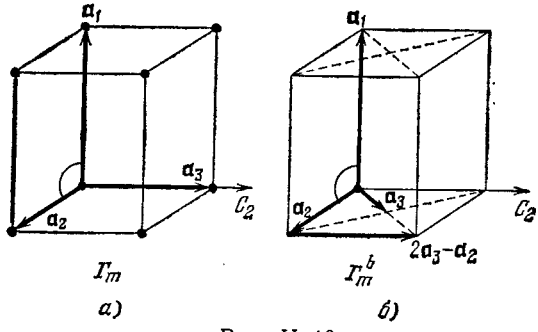


Рис. II.12.

вектор $\boldsymbol{a}_2$ на $\boldsymbol{a}_1+\boldsymbol{a}_2$. Существенно различными простыми решетками являются случаи 1) и один из трех остальных (например, 4)). Этим случаям, как это следует из (4.7), соответствуют следующие взаимные ориентации основных векторов (мы опускаем штрих у $\boldsymbol{a}_3'$):

$$1)\ \boldsymbol{a}_1,\ \boldsymbol{a}_2\perp\boldsymbol{a}_3,\quad 2)\ \boldsymbol{a}_1,\ \boldsymbol{a}_2\perp 2\boldsymbol{a}_3-\boldsymbol{a}_2.\qquad(4.8)$$

Решетки Браве, соответствующие случаям (4.8) (рис. II.12), называются *простой моноклинной* (Γ_m) и *моноклинной с центрированными основаниями* (Γ_m^b). Мы видим, что для типа Γ_m^b центру нижнего основания соответствует вектор $\boldsymbol{a}_3$, а центру верхнего основания—вектор $\boldsymbol{a}_3+\boldsymbol{a}_1$. Центры остальных четырех граней рассматриваемого параллелепипеда не совпадают с концами векторов трансляционной группы.

Заметим, что у типа Γ_m шесть векторов $\pm\boldsymbol{a}_1$, $\pm\boldsymbol{a}_2$, $\pm\boldsymbol{a}_3$ инвариантны относительно преобразования группы $\boldsymbol{C}_{2h}$. У типа Γ_m^b

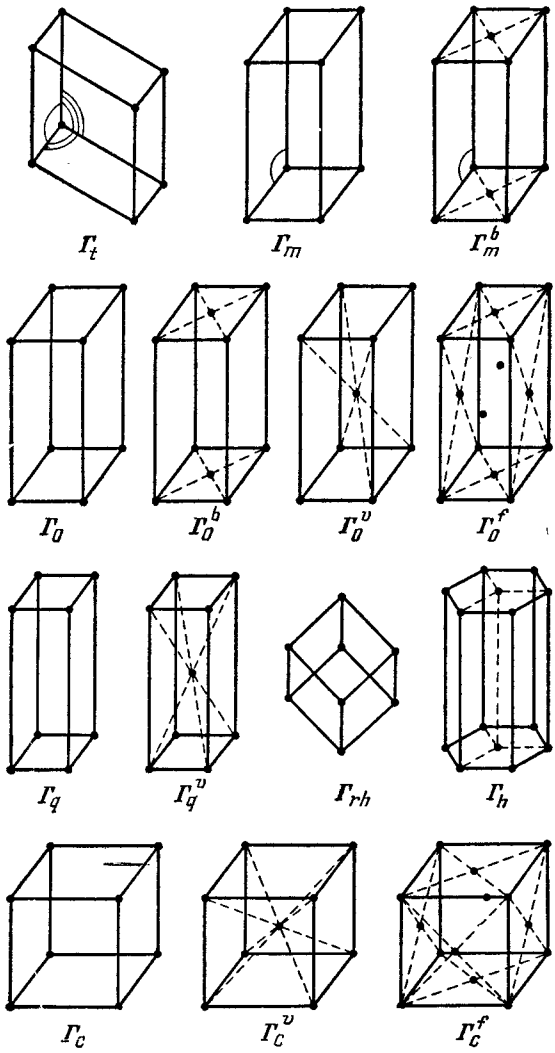


Рис. II.13.

инвариантными относительно этих преобразований являются векторы $\pm\boldsymbol{a}_1$, $\pm\boldsymbol{a}_2$ и $\pm(2\boldsymbol{a}_3-\boldsymbol{a}_2)$.

Исследуя аналогично пять остальных сингоний, можно получить все 14 решеток Браве, изображенных на рис. II.13.

Ромбической или ортогональной сингонии с симметрией D_{2h} соответствует четыре решетки Γ_0, Γ_0^b, Γ_0^v и Γ_0^f. Их элементарными ячейками являются прямоугольные параллелепипеды с тремя неравными ребрами. Типы решеток Γ_0 и Γ_0^b совершенно подобны решеткам Γ_m и Γ_m^b. Решетки Γ_0^v и Γ_0^f называются *ромбической объемноцентрированной* и *ромбической гранецентрированной*.

Тетрагональной сингонии с симметрией D_{4h} соответствуют решетки Γ_q и Γ_q^v. Они отличаются от ромбических решеток тем, что в их основаниях лежат не прямоугольники, а квадраты. Существуют два типа тетрагональных решеток: *простая* Γ_q и *объемноцентрированная* Γ_q^v. Центрирование верхнего и нижнего оснований у ячейки Γ_q не приводит к новому типу решетки, так как мы получим такую же простую тетрагональную структуру, но со стороной квадрата, в $\sqrt{2}$ раз меньшей.

Ромбоэдрической сингонии с симметрией D_{3d} соответствуют одна решетка Браве с $\alpha_{12}=\alpha_{23}=\alpha_{13}=60°$ и $a_1=a_2=a_3$ (обозначения см. на рис. I.4). Эта решетка может быть получена путем растяжения куба вдоль объемной диагонали. Она обозначена на рис. II.13 как Γ_{rh}.

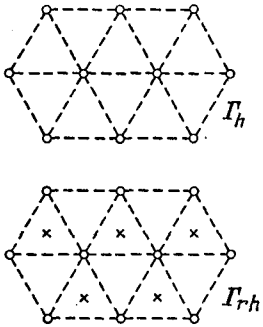


Рис. II.14.

Гексагональной сингонии с симметрией D_{6h} соответствует одна решетка Γ_h, в которой трансляционные векторы простой решетки могут быть направлены по трем ребрам шестигранной призмы, сходящимся в одной вершине.

Для дальнейшего полезно указать на особенности гексагональной и ромбоэдрической решеток. В обеих решетках атомы располагаются в плоскостях, перпендикулярных соответственно к осям 6-го и 3-го порядков. В обоих случаях атомы в плоскостях располагаются в вершинах равносторонних треугольников. Однако в то время, как в гексагональной решетке Γ_h атомы во всех плоскостях расположены друг над другом (поэтому гексагональная решетка обладает осью симметрии C_6; рис. II.14), в ромбоэдрической решетке Γ_{rh} атомы соседнего нижнего слоя расположены под центрами треугольников верхнего слоя (× и ○ на рис. II.14).

Наконец, как мы уже отметили, кубической сингонии с симметрией O_h соответствуют три решетки Браве: простая кубическая Γ_c, объемноцентрированная Γ_c^v и гранецентрированная Γ_c^f (рис. II.13). В гл. I, § 1, п. 3 мы рассмотрели вопрос, как должны быть выбраны основные векторы в объемноцентрированном и гранецентрированном кубах, для того чтобы примитивные ячейки содержали один атом.

§ 5. Кристаллические классы. Пространственные группы

1. В предыдущем параграфе мы рассмотрели свойства симметрии простых решеток. Мы пришли к заключению, что существуют только семь кристаллических систем (сингоний), которым может удовлетворять симметрия простых решеток, и 14 типов простых решеток (решеток Браве). Однако все химические соединения, а также некоторые простые вещества кристаллизуются, образуя сложные решетки. В этом случае примитивная ячейка кристалла содержит больше одного атома, в результате чего симметрия кристалла может только понизиться. Простейший пример этого мы уже указывали выше. Плоская сложная решетка на рис. I.2, *б* не имеет центра инверсии, которым обладает простая решетка, получающаяся при удалении атома, расположенного внутри ячейки.

Рассмотрим с этой точки зрения сложную решетку алмаза (Ge, Si), изображенную на рис. I.12. Как известно, эта решетка может быть получена в результате смещения одной гранецентрированной кубической решетки относительно другой вдоль объемной диагонали AB (рис. I.12) на $^1/_4$ ее длины. Кубическую ячейку алмаза можно представить себе состоящей из восьми малых кубов («кубиков»), пронумерованных на рис. I.12: *I*, *II*, ..., *VIII*. При описанном выше смещении (на $^1/_4$ $\overline{AB}$) ●-атомы *1'*, *2'*, *3'* и *4'* располагаются в центрах кубиков *I*, *III*, *VI*, *VIII*. Выбирая примитивную ячейку так, как это показано на рис. I.6, *а*, убедимся, что она содержит два атома.

Покажам, что симметрия сложной решетки алмаза ниже симметрии простой решетки гранецентрированного куба. Если поместить начало координат в один из узлов решетки алмаза, то она не инвариантна относительно инверсии. Оси C_4 гранецентрированного куба (рис. II.10) превращаются в решетке алмаза в оси C_2; в самом деле, при вращении вокруг вертикальной оси, проходящей через центр грани куба на угол $\pi/2$, атомы *1'* и *2'*, а также *3'* и *4'* не самосовмещаются, что имеет место только при повороте на угол π. Оси C_3, совпадающие с объемными диагоналями куба, являются и осями симметрии решетки алмаза; при вращении на угол $2\pi/3$ вокруг диагонали $\overline{AB}$ происходят следующие совпадения атомов: $2' \to 4'$, $3' \to 2'$, $4' \to 3'$. Наконец, плоскости отражения, проходящие через ось C_2 и две оси C_3 (рис. II.8, *в*) являются плоскостями симметрии решетки алмаза. Если провести плоскость отражения через вершины куба $ABCD$, то атомы *1'*, *2'* будут лежать в этой плоскости, а атомы *3'*, *4'* будут расположены относительно нее симметрично, поэтому плоскость эта будет плоскостью симметрии алмазной решетки. Таким образом, решетка алмаза обладает симметрией точечной группы T_d, являющейся подгруппой группы O_h.

Очевидно, что во всех случаях точечная группа сложного кристалла, получающаяся в результате заполнения атомами примитивной ячейки простой решетки, будет подгруппой сингонии простой решетки (решетки Браве).

Полный набор точечных групп $\mathscr{F}$ кристаллов, называемых *кристаллическими классами*, совпадает поэтому с набором всех возможных подгрупп, содержащихся в семи сингониях простых решеток Браве.

Выпишем все возможные подгруппы, входящие в семь сингоний:

Сингонии или кристаллические системы	*Кристаллические классы (подгруппы сингоний)*
1) триклинная $S_2 = C_i$:	E, $C_i = \{E, J\}$;
2) моноклинная C_{2h}:	$C_{1h} = C_s = \{E, \sigma_h\}$, C_2, C_{2h};
3) ромбическая D_{2h}:	C_{2v}, D_2, D_{2h};
4) тетрагональная D_{4h}:	C_4, C_{4v}, C_{4h}, S_4, D_{2d}, D_4, D_{4h}; (5.1)
5) ромбоэдрическая D_{3d}:	C_3, S_6, C_{3v}, D_3, D_{3d};
6) гексагональная D_{6h}:	C_6, C_{3h}, C_{6h}, C_{6v}, D_{3h}, D_6, D_{6h};
7) кубическая O_h:	T, T_h, T_d, O, O_h.

Итого: 32 кристаллических класса.

При определении кристаллических классов обычно оказывается, что каждый из них является подгруппой не одной сингонии, а нескольких. Во всех этих случаях кристаллический класс относят к системе (сингонии) наинизшей симметрии. Например, кристаллический класс $C_i = \{E, J\}$ является подгруппой всех сингоний, так как все простые решетки обладают центром инверсии J, но мы относим C_i к триклинной системе, не обладающей никакими другими элементами симметрии. При определении, к примеру, классов ромбической сингонии D_{2h} учитываем, что группа D_{2h} состоит из восьми элементов: E, C_2, $2C_2'$ (горизонтальные оси второго порядка), $2\sigma_v$, σ_h, $C_2\sigma_h = J$; подгруппы E, C_i, $C_{1h} = \{E, \sigma_h\}$ и C_{2h} должны быть отнесены к триклинной и моноклинной сингониям, поэтому к ромбической сингонии мы относим только кристаллические классы C_{2v}, D_2 и D_{2h}.

Условие, согласно которому мы относим класс к сингонии наинизшей симметрии, находит себе оправдание в физических соображениях. Маловероятно, чтобы атомы кристалла, относящиеся к определенному кристаллическому классу, образовали решетку Браве более симметричную, чем это необходимо для осуществления этого класса (можно думать, что это связано с минимумом термодинамического потенциала кристалла). Если бы подобная маловероятная ситуация осуществилась, то кристалл находился бы в метастабильном состоянии и малое возму-

щение (например, нагревание) перевело бы его в равновесное состояние с менее симметричной решеткой Браве. Если, например, кристалл класса D_4, который может реализоваться в тетрагональной системе, кристаллизовался бы в более симметричной кубической системе, то уже незначительное воздействие было бы способно удлинить или укоротить одно из ребер кубической ячейки, превратив ее в прямую призму с квадратным основанием.

Из этого примера видно, что для осуществления принципа кристаллизации в решетке Браве наинизшей симметрии необходима возможность непрерывного перехода посредством бесконечно малой деформации от решетки Браве высшей симметрии к низшей.

Сравнивая между собой решетки Браве разной симметрии, можно убедиться в том, что имеется одно исключение, когда такое непрерывное превращение невозможно. Именно, гексагональная решетка никакой бесконечно малой деформацией не может быть переведена в ромбоэдрическую решетку с более низкой симметрией. В самом деле, атомы соседних слоев ромбоэдрической решетки сдвинуты друг относительно друга на конечную величину, в то время как в гексагональной решетке они расположены друг под другом (см. рис. II.14). В этом смысле отнесение всех кристаллических классов к ромбоэдрической сингонии является условным, так как все они, являясь подгруппами группы D_{6h}, могут осуществляться в гексагональной структуре.

Таким образом, мы пришли к заключению, что *любой* сложный кристалл может быть описан определенной сингонией, решеткой Браве и кристаллическим классом.

Очевидно, что симметрия кристаллического класса, связанная с определенными поворотами, отражениями и инверсией, определяет физически эквивалентные направления в кристалле. Поскольку эта симметрия не связана с дискретными трансляциями, можно сказать, что она определяет макроскопическую симметрию анизотропного континуума. Такая симметрия определяет такие физические явления, как, например, распространение света в кристалле, его тепловое расширение и его механические свойства при действии внешних сил.

2. Мы изучили связь, которая существует между типами простых решеток (Браве) и их точечной симметрией (сингониями). Мы исследовали симметрию направлений в кристаллах (классы), определяющую их макроскопические свойства.

Перейдем к изучению полной симметрии кристаллов, т. е. к преобразованиям, приводящим к совпадению всех атомов одного сорта. Совокупность таких преобразований над кристаллом образует группу, называемую *пространственной группой*.

Эта симметрия кристалла, зависящая от расположения всех его атомов, которую естественно назвать микроскопической, определяет, например, рассеяние рентгеновских лучей кристаллом.

Всякий кристалл инвариантен (если отвлечься от наличия у него поверхности и дефектов) относительно трансляций с некоторыми основными векторами $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$.

Группа трансляций $\mathcal{T}$:

$$\boldsymbol{a} \equiv \boldsymbol{a}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3, \tag{5.2}$$

где n_1, n_2, n_3 — целые числа (включая нуль), есть инвариантная абелева подгруппа пространственной группы.

Наряду с трансляциями $\boldsymbol{a}_n$ пространственная группа, в общем случае, включает в себя различные преобразования точечной симметрии: вращения вокруг простых и зеркально-поворотных осей 2-го, 3-го, 4-го и 6-го порядков, отражения в плоскостях и инверсию (которая совпадает с зеркально-поворотной осью второго порядка S_2). Однако наряду с такими преобразованиями

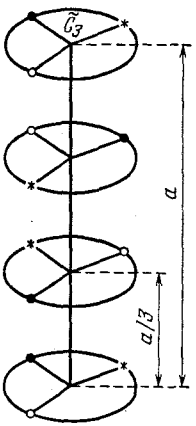


Рис. II.15.

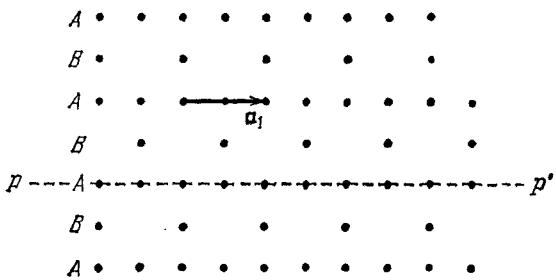


Рис. II.16.

точечной симметрии пространственная группа (т. е. кристалл) может обладать еще следующими элементами симметрии: *винтовой осью* и *плоскостью скольжения*.

Винтовой осью называется преобразование симметрии кристалла, состоящее из двух последовательных операций (взятых в произвольном порядке): вращения вокруг оси на угол $2\pi/n$ с последующим смещением вдоль оси на некоторую целую часть вектора решетки, называемым *несобственной трансляцией*.

На рис. II.15 изображена винтовая ось 3-го порядка $\tilde{C}_3$. Три атома разного сорта ●, ○, * расположены по окружности, перпендикулярной к оси, на равных расстояниях друг от друга. Собственной трансляции вдоль оси соответствует смещение на расстояние a. Однако самосовпадение имеет место и при повороте вокруг оси на угол $\alpha = 2\pi/3 = 120°$ с последующей несобственной трансляцией на величину $a/3$. Поскольку вращение и

поступательное движение соответствуют правому винту, ось $\tilde{C}_3$ называется *правовинтовой осью* 3-го порядка.

Рассмотрим еще один элемент симметрии решетки, связанный с несобственной трансляцией, — плоскость скольжения. На рис. II.16 изображена плоская (для наглядности) решетка, обладающая плоскостью скольжения PP'. При отражении в плоскости PP' решетка не совмещается сама с собой (слои A совпадают, слои B — нет). Если же после отражения сместить решетку в направлении PP' на $\boldsymbol{a}_1/2$, где $\boldsymbol{a}_1$ — основной вектор в направлении PP', то решетка совпадет сама с собой (можно вначале сдвинуть решетку на $\boldsymbol{a}_1/2$, а затем отразить ее в плоскости PP'). Можно показать, что в кристалле не могут существовать другие элементы симметрии, связанные с несобственной трансляцией, кроме двух рассмотренных выше.

Рассмотрим решетку алмаза (рис. I.12) с точки зрения существования у нее *несобственных элементов симметрии*, т. е. элементов симметрии, связанных с несобственными трансляциями. Будем называть гранецентрированную кубическую решетку, состоящую из ○-атомов, основной, а из ●-атомов сдвинутой.

В п. 1 настоящего параграфа мы показали, что решетка алмаза инвариантна относительно преобразований точечной группы $\boldsymbol{T}_d$. Исследуем для нее операции $J\boldsymbol{T}_d = \boldsymbol{T}_d J$ (где J — инверсия). Так как операции группы $\boldsymbol{T}_d$ оставляют решетку алмаза инвариантной, достаточно исследовать действие инверсии J. При применении инверсии к ○-атомам основной решетки (выберем атом A основной решетки за начало координат) мы воспроизведем основную решетку, которая после несобственной трансляции на $^1/_4\ \overline{AB}$ совпадает со сдвинутой решеткой ●-атомов. При применении же инверсии к ●-атомам сдвинутой решетки (при том же начале координат в A) они попадут на «пустые места», но после несобственной трансляции $^1/_4\ \overline{AB}$ они совпадут с ○-атомами основной решетки. Таким образом, операции $J\boldsymbol{T}_d$, сопровождаемые несобственной трансляцией $^1/_4\ \overline{AB}$, также являются операциями симметрии решетки алмаза.

Читатель может быть смущен тем, что, как выше утверждалось, существуют только две несобственные операции симметрии кристалла — винтовая ось и плоскость скольжения, в то время как в алмазной решетке появляется, как будто бы новый несобственный элемент симметрии, связанный с инверсией.

Следует, однако, сказать, что выделение несобственных элементов симметрии в некотором отношении условно. Если бы мы поместили начало координат не в атоме A, а посредине между ним и ●-атомом $1'$, то решетка алмаза была бы инвариантна относительно инверсии (не сопровождаемой несобственной

трансляцией). Однако при этом появились бы в решетке алмаза плоскости скольжения [1]).

В заключение заметим, что сурьмянистый индий (InSb), кристаллизующийся в алмазоподобной решетке, не обладает несобственными элементами симметрии, так как в этом случае ○-атомы и ●-атомы (рис. I.12) разного сорта (In и Sb). Симметрия InSb есть точечная группа T_d.

3. Все возможные пространственные группы кристаллов получаются при согласовании всех решеток Браве со всеми возможными осями вращательной симметрии, плоскостями отражения, винтовыми осями и плоскостями скольжения. Для того чтобы представить себе особенности такого согласования, введем оператор (элемент) пространственной группы $\{R\,|\,\alpha(R)+\boldsymbol{a}_n\}$, где R—элемент точечной группы кристаллического класса $\mathcal{F}$ (поворот, зеркальный поворот, отражение в плоскости, инверсия), $\alpha(R)$—несобственная трансляция (обязательно рациональная часть вектора решетки), соответствующая элементу R, и $\boldsymbol{a}_n$—вектор решетки; действие этого оператора на радиус-вектор $\boldsymbol{r}$ определим равенством [2])

$$\{R\,|\,\alpha(R)+\boldsymbol{a}_n\}\,\boldsymbol{r}=\boldsymbol{R}\boldsymbol{r}+\alpha(R)+\boldsymbol{a}_n. \tag{5.3}$$

Здесь $\boldsymbol{R}$ в правой части равенства—матрица ортогонального преобразования вращения, инверсии или отражения R. Для тех элементов $\boldsymbol{R}$, для которых несобственная трансляция равна нулю: $\alpha(R)=0$. Для того чтобы оператор

$$g=\{R\,|\,\alpha(R)+\boldsymbol{a}_n\} \tag{5.4}$$

являлся элементом пространственной группы кристалла, необходимо, чтобы множество (5.4) содержало единичный и обратные элементы и чтобы произведение двух элементов множества (5.4) принадлежало той же совокупности.

Очевидно, что единичный элемент множества (5.4) равен $\{E\,|\,0\}$, где E—единичный элемент точечной группы $\mathcal{F}$, в самом деле,

$$\{E\,|\,0\}\,\boldsymbol{r}=E\boldsymbol{r}=\boldsymbol{r}. \tag{5.5}$$

Обратный элемент (5.4) равен

$$g^{-1}=\{R\,|\,\alpha(R)+\boldsymbol{a}_n\}^{-1}=\{R^{-1}\,|\,-\boldsymbol{R}^{-1}\alpha(R)-\boldsymbol{R}^{-1}\boldsymbol{a}_n\}, \tag{5.6}$$

в самом деле,

$$\begin{aligned}g^{-1}g\boldsymbol{r}&=\{R^{-1}\,|\,-\boldsymbol{R}^{-1}\alpha(R)-\boldsymbol{R}^{-1}\boldsymbol{a}_n\}\,\{\boldsymbol{R}\boldsymbol{r}+\alpha(R)+\boldsymbol{a}_n\}=\\&=\boldsymbol{R}^{-1}\boldsymbol{R}\boldsymbol{r}+\boldsymbol{R}^{-1}\alpha(R)+\boldsymbol{R}^{-1}\boldsymbol{a}_n-\boldsymbol{R}^{-1}\alpha(R)-\boldsymbol{R}^{-1}\boldsymbol{a}_n=\boldsymbol{r}=\\&=\{E\,|\,0\}\,\boldsymbol{r}.\end{aligned} \tag{5.7}$$

[1]) Джонс Г. Теория зон Бриллюэна и электронных состояний в кристаллах. — М.: Мир. 1968, § 34.

[2]) Мы пользуемся распространенным обозначением оператора (элемента) пространственной группы, где сперва действует на $\boldsymbol{r}$ оператор R, стоящий слева от вертикальной черты, а затем оператор $\alpha(R)+\boldsymbol{a}_n$ (смещение), стоящий справа от черты.

Наконец, покажем, при каких условиях произведение двух элементов множества (5.4) принадлежит тому же множеству. Если R_1 и R_2—два элемента точечной группы кристалла $\mathcal{F}$, то

$$\{R_2|\alpha(R_2)+a_m\}\{R_1|\alpha(R_1)+a_n\}r=$$
$$=\{R_2|\alpha(R_2)+a_m\}(R_1r+\alpha(R_1)+a_n\}=$$
$$=R_2R_1r+R_2\alpha(R_1)+R_2a_n+\alpha(R_2)+a_m. \quad (5.8)$$

Первое слагаемое правой части (5.8) имеет структуру первого слагаемого правой части (5.3), так как $R_2R_1=R=$ элементу точечной группы $\mathcal{F}$. Жестким условием для элементов пространственной группы (5.4) является требование, чтобы остальные четыре слагаемых в (5.8) имели структуру (5.3), т. е.

$$R_2\alpha(R_1)+R_2a_n+\alpha(R_2)+a_m=\alpha(R)+a_p, \quad (5.9)$$

где $\alpha(R)$—несобственная трансляция, соответствующая элементу $R=R_2R_1$, а a_p—какой-либо вектор решетки. Ясно, что (5.9) может быть выполнено только при определенной согласованности операций $\alpha(R_1)$, несобственных трансляций $\alpha(R_1)$, $\alpha(R_2)$, $\alpha(R_2R_1)$ и основных векторов решетки a_1, a_2, a_3. Тщательный анализ, проведенный в 1891 г., независимо, русским кристаллографом Е. С. Федоровым и немецким А. Шёнфлисом, показал, что всего имеется 230 пространственных групп [1]).

Пространственные группы, у которых $\alpha(R)=0$ для всех R, называются *простыми* или *симморфными*; если же хотя бы для одного R, $\alpha(R)\neq 0$, группа называется *несимморфной*.

Для симморфных групп условие (5.9) упрощается, приобретая вид

$$R_2a_n+a_m=a_p. \quad (5.10)$$

Если R_2—операция вращения вокруг оси, то требование (5.10) привело нас к возможности существования осей только 2-го, 3-го, 4-го и 6-го порядков (§ 4, п. 1).

Можно получить представление о числе симморфных пространственных групп, если принять во внимание, что каждый кристаллический класс, комбинируя с возможной решеткой Браве, образует пространственную группу. Например, ромбический сингонии D_{2h} соответствуют три класса (см. (5.1)) и четыре решетки Браве Γ_0, Γ_0^b, Γ_0^v, Γ_0^f (рис. II.13), поэтому мы получим для этого случая по крайней мере двенадцать пространственных симморфных групп. Учитывая, что классы ромбоэдрической сингонии могут содержаться и в гексагональной, получим для всех 14 решеток Браве и 32 классов (5.1) 61 пространственную

[1]) Федоров Е. С. Симметрия и структура кристаллов.— М., 1949; Schönflies A. Theorie der Kristallstrukturen.— Berlin, 1923.

Таблица II.2

Сингонии или кристаллические системы	Решетки Браве	Точечные группы	Число пространственных групп	Полное число пространственных групп
Триклинная	Триклинная	E C_i	1 1	2
Моноклинная	Простая моноклинная, двухгранецентрированная моноклинная	C_{1h} C_2 C_{2h}	4 3 6	13
Ромбическая или ортогональная	Простая ромбическая, двухгранецентрированная ромбическая, центрированная ромбическая, объемноцентрированная ромбическая	C_{2v} D_2 D_{2h}	22 9 28	59
Тетрагональная или квадратная	Простая тетрагональная, объемноцентрированная тетрагональная	C_4 C_{4v} C_{4h} S_4 D_{2d} D_4 D_{4h}	6 12 6 2 12 10 20	68
Ромбоэдрическая или тригональная	Тригональная	C_3 S_6 C_{3v} D_3 D_{3d}	4 2 6 7 6	25
Гексагональная	Гексагональная	C_6 C_{3h} C_{6h} C_{6v} D_{3h} D_6 D_{6h}	6 1 2 4 4 6 4	27
Кубическая	Простая кубическая, гранецентрированная кубическая, объемноцентрированная кубическая	T T_h T_d O O_h	5 7 6 8 10	36
Итого	14	32	230	230

симморфную группу. Если еще учесть, что в некоторых случаях элементы симметрии кристаллического класса могут быть по-разному расположены относительно трансляционных векторов (например, двойные горизонтальные оси в классе $\boldsymbol{D_3}$ гексагональной структуры можно направить к вершинам шестиугольника и к серединам его сторон), то полное число симморфных групп возрастает до 73.

Винтовые оси, плоскости скольжения представляют широкие дополнительные возможности конструирования пространственных групп; число несимморфных групп оказывается равным 157, так что общее число пространственных групп равно $73+157=230$.

В табл. II.2 дано распределение всех пространственных групп по сингониям и кристаллическим классам.

Представим элемент пространственной группы (5.4) в виде произведения двух операторов

$$g=\left\{R\left|\boldsymbol{\alpha}(R)+\boldsymbol{a}_n\right.\right\}=\left\{E\left|\boldsymbol{a}_n\right.\right\}\left\{R\left|\boldsymbol{\alpha}(R)\right.\right\}, \qquad (5.11)$$

где E — единичный элемент точечной группы $\mathcal{F}$ с элементами R. Хотя преобразования $\left\{R\left|\boldsymbol{\alpha}(R)\right.\right\}$ являются элементами симметрии кристалла, они не образуют группу. В самом деле, их произведение может дать вектор решетки $\boldsymbol{a}$, принадлежащей к группе трансляций $\mathcal{T}$.

Например, применяя три раза преобразование, соответствующее винтовой оси на рис. II.15, получим смещение вдоль оси на величину основного вектора $\boldsymbol{a}$. Аналогично применяя два раза оператор $\left\{R\left|\frac{\boldsymbol{a}_1}{2}\right.\right\}$, где R — отражение в плоскости PP' (рис. II.16), получим трансляцию на основной вектор $\boldsymbol{a}_1$.

В то же время совокупность всех ортогональных преобразований R (в том числе и тех, для которых $\boldsymbol{\alpha}(R)\neq 0$) образует точечную группу $\mathcal{F}$, определяющую класс кристалла. В самом деле, из (5.8) видно, что если R_1 и R_2 — ортогональные преобразования симметрии кристалла (с несобственными трансляциями, равными или не равными нулю), то $R_2R_1=R$ — тоже ортогональное преобразование симметрии кристалла (с $\boldsymbol{\alpha}(R)=0$ или $\boldsymbol{\alpha}(R)\neq 0$).

Таким образом, для того чтобы определить класс кристалла, надо учесть все оси и плоскости симметрии, заменив при этом винтовые оси и плоскости скольжения на равнозначные простые оси и плоскости отражения.

§ 6. Неприводимые представления групп и теория характеров

1. Применение теории групп в физике и, в частности, к твердому телу, основано на теории неприводимых представлений и теории характеров [1]).

Рассмотрим некоторую группу $\boldsymbol{G}$, состоящую из h элементов: $\boldsymbol{A} \equiv E, B, C, \ldots, S, T, \ldots$ Для определенности представим себе, что это точечная группа симметрии, каждому элементу которой соответствует некоторое преобразование координат тела x, y, z. Возьмем некоторую произвольную однозначную функцию $\psi_1(x, y, z) = \psi_1(\boldsymbol{x})$ и применим к ней оператор $\hat{P}_R$ (R — один из элементов группы), определенный следующим образом:

$$\hat{P}_R \psi_1(x, y, z) = \psi_1(\boldsymbol{R}^{-1}\boldsymbol{x}) = \Phi_R(\boldsymbol{x}). \tag{6.1}$$

Здесь R^{-1} — элемент, обратный R, $\boldsymbol{R}^{-1}\boldsymbol{x}$ — ортогональное преобразование координат, соответствующее элементу R^{-1}. На первый взгляд было бы естественнее определить оператор $\hat{P}_R$ так:

$$\hat{P}'_R \psi_1(\boldsymbol{x}) = \psi_1(\boldsymbol{R}\boldsymbol{x}). \tag{6.1а}$$

Мы покажем, однако, что только при определении (6.1) операторы $\hat{P}_R$ образуют группу, изоморфную с группой $\boldsymbol{G}$, т. е.

$$\hat{P}_S \hat{P}_R = \hat{P}_{SR}. \tag{6.1б}$$

В самом деле, из определения оператора $\hat{P}_R$ (6.1) следует

$$\hat{P}_S \hat{P}_R \psi_1(\boldsymbol{x}) = \hat{P}_S \Phi_R(\boldsymbol{x}) = \Phi_R(\boldsymbol{S}^{-1}\boldsymbol{x}) = \psi_1(\boldsymbol{R}^{-1}\boldsymbol{S}^{-1}\boldsymbol{x}) = \\ = \psi_1[(\boldsymbol{S}\boldsymbol{R})^{-1}x] = \hat{P}_{SR}\psi_1(\boldsymbol{x}),$$

откуда прямо следует (6.1б). При определении же (6.1а) мы получили бы $\hat{P}'_S \hat{P}'_R = \hat{P}'_{RS}$, что не всегда удобно [2]). Взяв в (6.1) вместо R все элементы группы $\boldsymbol{G}$, мы, в общем случае, не получим h линейно независимых функций $\Phi_R(\boldsymbol{x})$; число линейно независимых функций r будет, вообще говоря, меньше или равно h, т. е. $r \leqslant h$; обозначим их через $\psi_1(\boldsymbol{x})$, $\psi_2(\boldsymbol{x})$, ..., $\psi_r(\boldsymbol{x})$, где $\psi_1(\boldsymbol{x})$ — одна из этих функций (так как R может равняться E). Посредством соответствующего линейного преобразования можно всегда функции $\psi_i(\boldsymbol{x})$ ($i = 1, 2, \ldots, r$) сделать ортонормированными, что мы и будем предполагать в дальнейшем. Функции $\psi_i(\boldsymbol{x})$ ($i = 1, 2, \ldots, r$) называются: *базисные функции*, *базисный вектор* или просто *базис*. При другом выборе функции $\psi_1(\boldsymbol{x})$, число базис-

[1]) Необходимые для понимания этого параграфа сведения о матрицах и их свойствах содержатся в Приложении 3.

[2]) Некоторые авторы, например в книге Solid State Theory./P. T. Landsberg.— London, 1969, с. 157, предпочитают все-таки определение (6.1а).

ных функций r, вообще говоря, меняется. Применение $\hat{P}_S$ к одной из базисных функций, например к $\psi_i(\boldsymbol{x})$, выражается через линейную суперпозицию базисных функций (это следует из (6.1), если применить к обеим частям равенства $\hat{P}_S$ и учесть, что $\Phi_R(\boldsymbol{x})$ — линейная суперпозиция функций $\psi_i(\boldsymbol{x})$, т. е.

$$\hat{P}_S \psi_i = \sum_{k=1}^{r} \Gamma(S)_{ki} \psi_k, \tag{6.2}$$

где $\Gamma(S)_{ki}$ — элемент матрицы r-го ранга $\Gamma(S)$ [1]).

Умножая обе части равенства (6.2) слева на $\psi_l^*(x)$, интегрируя по $d\boldsymbol{x} = dx\,dy\,dz$ и используя ортонормированность базисных функций, получим

$$\int \psi_l^*(\boldsymbol{x})\, \hat{P}_S\, \psi_i(\boldsymbol{x})\, d\boldsymbol{x} = \sum_{k=1}^{r} \Gamma(S)_{ki} \int \psi_l^*(\boldsymbol{x})\, \psi_k(\boldsymbol{x})\, d\boldsymbol{x} =$$

$$= \sum_{k=1}^{r} \Gamma(S)_{ki}\, \delta_{lk} = \Gamma(S)_{li}. \tag{6.3}$$

Таким образом, элементы матрицы $\Gamma(S)$ в (6.2) — матричные элементы оператора $\hat{P}_S$, взятые на базисных функциях $\psi_i(\boldsymbol{x})$. Если к обеим частям равенства (6.2) применить оператор $\hat{P}_T$, то

$$\hat{P}_T \hat{P}_S \psi_i = \hat{P}_{TS} \psi_i = \sum_l \Gamma(TS)_{li} \psi_l = \sum_k \Gamma(S)_{ki} \hat{P}_T \psi_k =$$

$$= \sum_k \Gamma(S)_{ki} \sum_l \Gamma(T)_{lk} \psi_l = \sum_l \sum_k \Gamma(T)_{lk}\, \Gamma(S)_{ki}\, \psi_l =$$

$$= \sum_l [\Gamma(T)\, \Gamma(S)]_{li}\, \psi_l, \tag{6.4}$$

где мы в конце воспользовались законом умножения матриц. Здесь первая строка получена от применения $\hat{P}_T$ к левой части (6.2), а вторая строка — к правой части (6.2).

Сравнивая третье звено с последним, получим

$$\Gamma(TS)_{li} = [\Gamma(T)\, \Gamma(S)]_{li}, \tag{6.5}$$

или

$$\Gamma(TS) = \Gamma(T)\, \Gamma(S). \tag{6.5а}$$

Мы видим, что матрица, соответствующая произведению TS, равна произведению матриц для T и S. Матрицы $\Gamma(S)$ в (6.2), где S — любой элемент группы, перемножающиеся по той же таблице умножения, что и элементы группы, называются *представлением группы*. Если все матрицы представления различны, то оно назы-

[1]) Такое сопряжение номера базисной функции с первым индексом матрицы в сумме (6.2) целесообразно в смысле сопоставления матриц элементам группы (см. ниже).

вается *точным*, в противном случае оно называется *неточным*. В первом случае матричное представление образует группу, изоморфную исходной группе; во-втором — исходная группа гомоморфна матричному представлению.

Если воспользоваться определением оператора $\hat{P}'_R$ (6.1а), то вместо (6.5а) получим

$$\Gamma(TS) = \Gamma(S)\,\Gamma(T), \tag{6.5б}$$

т. е. матрицы представлений перемножаются в порядке, обратном элементам группы (это связано с тем, что, как отмечено выше, в случае (6.1а) $\hat{P}'_{TS} = \hat{P}'_S \hat{P}'_T$). Однако транспонированные матрицы представлений и при определении $\hat{P}'_R$ (6.1а) перемножаются «правильно» согласно (6.5а). В самом деле, из (6.5б) следует

$$\tilde{\Gamma}(TS) = \widetilde{\Gamma(S)\,\Gamma(T)} = \tilde{\Gamma}(T)\,\tilde{\Gamma}(S), \tag{6.5в}$$

где было использовано (П.3.30).

В следующем параграфе мы увидим, как естественно возникают представления и базисные функции при исследовании симметрии квантовомеханических систем.

Единичному элементу группы E соответствует единичная матрица

$$\Gamma(E) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \|\delta_{ik}\|, \tag{6.6}$$

где $\delta_{ik} = 1$ при $i = k$ и $\delta_{ik} = 0$ при $i \neq k$. В самом деле,

$$\Gamma(E)\,\Gamma(S) = \Gamma(S). \tag{6.7}$$

Действительно,

$$[\Gamma(E)\,\Gamma(S)]_{ik} = \sum_l \Gamma(E)_{il}\,\Gamma(S)_{lk} = \sum_l \delta_{il}\Gamma(S)_{lk} = \Gamma(S)_{ik},$$

что совпадает с (6.7) (т. е. мы показали, что матричный элемент произведения матриц в левой части (6.7) равен соответствующему матричному элементу матрицы $\Gamma(S)$).

Если $Q^{-1}Q = E$, т. е. Q^{-1} — элемент, обратный Q, то

$$\Gamma(Q^{-1}) = \Gamma(Q)^{-1}, \tag{6.8}$$

т. е. матрица обратного элемента есть обратная матрица прямого элемента. В самом деле, $\Gamma(Q^{-1})\,\Gamma(Q) = \Gamma(E)$; умножая обе части справа на обратную матрицу $\Gamma(Q)^{-1}$, получим (6.8). Из (6.8) следует, что представление группы может осуществляться только такими матрицами, которые имеют обратные, т. е. несингулярными матрицами (определитель таких матриц не должен равняться нулю).

Если все матрицы представления подвергнуть *преобразованию подобия* (см. Приложение 3, п. 4)

$$S^{-1}\Gamma(A)S=\Gamma(A)',\quad S^{-1}\Gamma(B)S=\Gamma(B)',\quad S^{-1}\Gamma(C)S=\Gamma(C)',\ \ldots, \tag{6.9}$$

где S — некоторая (несингулярная) матрица, то новые («штрихованные») матрицы тоже дают представления группы. В самом деле,

$$\Gamma(B)'\Gamma(C)'=S^{-1}\Gamma(B)SS^{-1}\Gamma(C)S=$$
$$=S^{-1}\Gamma(B)\Gamma(C)S=S^{-1}\Gamma(BC)S=\Gamma(BC)', \tag{6.10}$$

где во втором равенстве использовано, что $SS^{-1}=E$. Так как выбор матрицы S произволен, то очевидно, что бесконечное многообразие представлений (6.9) не может существенно отличаться друг от друга и давать разную информацию. Мы будем все такие представления, получаемые посредством преобразования подобия, называть *эквивалентными* и будем рассматривать их как одно представление. В связи с этим важно отметить, что шпур (след) матрицы не меняется при преобразовании подобия:

$$\mathrm{Sp}\,\Gamma(A)'=\mathrm{Sp}\,[S^{-1}\Gamma(A)S]=\mathrm{Sp}\,[\Gamma(A)SS^{-1}]=$$
$$=\mathrm{Sp}\,[\Gamma(A)E]=\mathrm{Sp}\,\Gamma(A), \tag{6.11}$$

где мы воспользовались тем, что шпур произведения двух матриц не зависит от порядка их умножения. Таким образом, матрицы всех эквивалентных представлений имеют для данного элемента группы один и тот же шпур.

Покажем, что если базисные функции ортонормированы, т. е. их скалярное произведение

$$(\psi_i,\ \psi_k)=\int\psi_i^*(\boldsymbol{x})\,\psi_k(\boldsymbol{x})\,d\boldsymbol{x}=\delta_{ik}, \tag{6.12}$$

то матрицы представления $\Gamma(R)$ унитарны.

В самом деле,

$$\delta_{ik}=(\psi_i,\ \psi_k)=(\hat{P}_R\psi_i,\ \hat{P}_R\psi_k)=\Big(\sum_r\Gamma(R)_{ri}\,\psi_r,\ \sum_s\Gamma(R)_{sk}\,\psi_s\Big)=$$
$$=\sum_{r,s}\Gamma^*(R)_{ri}\,\Gamma(R)_{sk}\,(\psi_r,\ \psi_s)=\sum_{r,s}\Gamma^*(R)_{ri}\,\Gamma(R)_{sk}\delta_{rs}=$$
$$=\sum_r\Gamma^*(R)_{ri}\,\Gamma(R)_{rk}=\sum_r\Gamma^+(R)_{ir}\,\Gamma(R)_{rk}=[\Gamma^+(R)\,\Gamma(R)]_{ik}.$$

При переходе в этой цепи равенств от второго звена к третьему мы воспользовались тем, что скалярное произведение $(\psi_i\psi_k)$ не меняется при ортогональном преобразовании $\hat{P}_R$ переменных интегрирования (якобиан преобразования равен единице); при переходе от пятого звена к шестому мы воспользовались ортонормированностью функций $\psi_i(\boldsymbol{x})$; наконец, при переходе от седьмого звена к восьмому — определением сопряженной матрицы.

Из сравнения первого звена с последним следует

$$\Gamma^{+}(R)\,\Gamma(R)=E, \tag{6.13}$$

т. е. матрицы $\Gamma(R)$ унитарны.

Нетрудно показать (на этом мы останавливаться не будем), что если старые базисные функции подвергнуть произвольному линейному преобразованию

$$\psi_i'=\sum_k S_{ki}\psi_k, \tag{6.14}$$

то новое («штрихованное») представление

$$\Gamma'(R)=\boldsymbol{S}^{-1}\Gamma(R)\,\boldsymbol{S}, \tag{6.15}$$

т. е. эквивалентно старому.

Легко убедиться (посредством прямого испытания), что группа равностороннего треугольника $\boldsymbol{D}_3$ (§ 2, табл. II.1) имеет три следующие (неэквивалентные) представления: 1) тривиальное единичное представление, когда всем элементам группы сопоставляется единица, группа $\boldsymbol{D}_3$ гомоморфна этому представлению размерности единица (таким представлением обладает любая группа); 2) представление размерности единица, в виде: $\Gamma(E)=1$, $\Gamma(A)=\Gamma(B)=\Gamma(C)=-1$, $\Gamma(D)=\Gamma(F)=1$, что легко проверить, используя табл. II.1; 3) изоморфное представление размерности два:

$$\Gamma(E)=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \Gamma(A)=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad \Gamma(B)=\begin{pmatrix}-1/2&\sqrt{3}/2\\\sqrt{3}/2&1/2\end{pmatrix}, \tag{6.16}$$

$$\Gamma(C)=\begin{pmatrix}-1/2&-\sqrt{3}/2\\-\sqrt{3}/2&1/2\end{pmatrix},\quad \Gamma(D)=\begin{pmatrix}-1/2&\sqrt{3}/2\\-\sqrt{3}/2&-1/2\end{pmatrix},$$

$$\Gamma(F)=\begin{pmatrix}-1/2&-\sqrt{3}/2\\\sqrt{3}/2&-1/2\end{pmatrix}.$$

В самом деле, например,

$$\Gamma(A)\,\Gamma(D)=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}-1/2&\sqrt{3}\\-\sqrt{3}/2&-1/2\end{pmatrix}=\begin{pmatrix}-1/2&\sqrt{3}/2\\\sqrt{3}/2&1/2\end{pmatrix}=\Gamma(B),$$

$$\Gamma(F)\,\Gamma(D)=\begin{pmatrix}-1/2&-\sqrt{3}/2\\\sqrt{3}/2&-1/2\end{pmatrix}\begin{pmatrix}-1/2&\sqrt{3}/2\\-\sqrt{3}/2&-1/2\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}=\Gamma(E),$$

в соответствии с законом умножения матриц и табл. II.1. Заметим, что все матрицы (6.16) унитарны. Так как все матрицы представления (6.16) различны, то оно называется точным.

Выбрав какую-либо произвольную несингулярную матрицу второго ранга $\boldsymbol{S}$ и используя (6.9), мы могли бы матричное представление (6.16) написать в совершенно ином внешнем, но по существу эквивалентном виде.

Двумерное представление (6.16) может быть построено, если воспользоваться базисными функциями:

$$\psi_1 = \sqrt{\frac{2}{\pi}}\, x \exp\left[-\frac{1}{2}(x^2+y^2)\right], \quad \psi_2 = \sqrt{\frac{2}{\pi}}\, y \exp\left[-\frac{1}{2}(x^2+y^2)\right].$$

Для определения матриц $\Gamma(S)_{ik}$, входящих в (6.2), подвергнем координаты x, y преобразованиям группы $\boldsymbol{D}_3$ (§ 2, п. 2)[1]); при этом мы можем вращать треугольник, считая неподвижными оси координат x, y (так определены операции A, B, C, ...), или вращать координатную систему (в обратном направлении), считая неподвижным треугольник (так в нашем случае это удобнее). Из рис. II.1 следует, что операциям симметрии группы $\boldsymbol{D}_3$ соответствуют преобразования (6.2)

$$\hat{P}_A\psi_1\colon\ \hat{A}^{-1}x = x,\ P_A\psi_2\colon\ \hat{A}^{-1}y = -y$$

или

$$\hat{P}_A\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}, \quad \Gamma(A) = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix};$$

$$\hat{P}_B\psi_1\colon\ \hat{B}^{-1}x = -\frac{1}{2}x + \frac{\sqrt{3}}{2}y, \quad \hat{P}_B\psi_2\colon\ \hat{B}^{-1}y = \frac{\sqrt{3}}{2}x + \frac{1}{2}y$$

или

$$\hat{P}_B\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}-1/2 & \sqrt{3}/2\\ \sqrt{3}/2 & 1/2\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}, \quad \Gamma(B) = \begin{pmatrix}-1/2 & \sqrt{3}/2\\ \sqrt{3}/2 & 1/2\end{pmatrix}.$$

Аналогично можно рассмотреть преобразования, соответствующие элементам группы C, D, F, т. е. построить все матрицы (6.16).

2. Представим себе, что нам даны два неэквивалентных представления, вообще говоря, разной размерности r и s:

$$\begin{aligned}&\Gamma_1(A),\ \Gamma_1(B),\ \Gamma_1(C),\ \ldots,\ \Gamma_1(P),\ \ldots,\\ &\Gamma_2(A),\ \Gamma_2(B),\ \Gamma_2(C),\ \ldots,\ \Gamma_2(P),\ \ldots\end{aligned} \tag{6.17}$$

Составим из матриц Γ_1 r-го ранга и матриц Γ_2 s-го ранга *блочные* или *квазидиагональные* матрицы $(r+s)$-го ранга вида

$$\begin{gathered}\Gamma(A) = \left(\begin{array}{c|c}\Gamma_1(A) & 0\\ \hline 0 & \Gamma_2(A)\end{array}\right), \quad \Gamma(B) = \left(\begin{array}{c|c}\Gamma_1(B) & 0\\ \hline 0 & \Gamma_2(B)\end{array}\right),\\ \Gamma(C) = \left(\begin{array}{c|c}\Gamma_1(C) & 0\\ \hline 0 & \Gamma_2(C)\end{array}\right) \text{ и т. д.,}\end{gathered} \tag{6.18}$$

где нули в правом верхнем углу заполняют прямоугольный блок шириной s столбцов и высотой r строк, а нули в левом нижнем углу — прямоугольный блок шириной r столбцов и высотой s строк. Покажем, что матрицы Γ (6.18) тоже дают представление

[1]) Заметим, что множитель $\sqrt{\frac{2}{\pi}}\exp\left[-\frac{1}{2}(x^2+y^2)\right]$ остается инвариантным при всех преобразованиях группы $\boldsymbol{D}_3$.

группы, т. е. если $BA=C$, то

$$\Gamma(B)\,\Gamma(A)=\Gamma(BA)=\Gamma(C).$$

В самом деле, из формулы, определяющей матричные элементы матрицы, равной произведению двух других матриц, следует:

$$[\Gamma(B)\,\Gamma(A)]_{ik}=\sum_l \left(\begin{array}{c|c} \Gamma_1(B) & 0 \\ \hline 0 & \Gamma_2(B) \end{array}\right)_{il} \left(\begin{array}{c|c} \Gamma_1(A) & 0 \\ \hline 0 & \Gamma_2(A) \end{array}\right)_{lk},$$

где сумма в правой части вычисляется по схеме

$$[\ \ ]_{ik}=i\,(*\ *\ *)\begin{pmatrix} * \\ * \\ * \end{pmatrix}k,$$

т. е. элементы i-й строки первой матрицы умножаются на элементы k-го столбца второй матрицы. Мы видим, что до тех пор пока i меняется от 1 до r, k тоже меняется только от 1 до r (так как большим значениям k соответствуют нули). Таким образом, левый верхний блок $\Gamma_1(B)$ умножается на левый верхний блок $\Gamma_1(A)$. Аналогичная ситуация имеет место для нижних правых блоков; поэтому

$$\Gamma(B)\,\Gamma(A)=\left(\begin{array}{c|c} \Gamma_1(B)\,\Gamma_1(A) & 0 \\ \hline 0 & \Gamma_2(B)\,\Gamma_2(A) \end{array}\right)=\left(\begin{array}{c|c} \Gamma_1(C) & 0 \\ \hline 0 & \Gamma_2(C) \end{array}\right)=\Gamma(C),$$

что и требовалось доказать.

Если мы подвергнем теперь матрицы Γ в представлении (6.18) какому-либо преобразованию подобия (посредством какой-либо матрицы $(r+s)$-го ранга), то матрицы Γ потеряют свой квазидиагональный (блочный) вид, хотя по-прежнему будут давать представление группы. Подвергнув это эквивалентное представление обратному преобразованию подобия, мы приведем его опять к квазидиагональному (блочному) виду.

Таким образом, мы приходим к заключению, что возможны такие случаи, когда, подвергнув все матрицы представления преобразованию подобия, мы приведем их к квазидиагональному виду (с одинаковой структурой блоков). В этом случае представление называется *приводимым*.

Весьма важным является утверждение, что в некоторых случаях матрицы представления никаким преобразованием подобия не могут быть приведены к квазидиагональному виду. Такие представления называются *неприводимыми*.

Существование неприводимых представлений в форме матриц второго или более высокого ранга представляется весьма естественным для неабелевых групп, так как в этом случае некоммутативности умножения элементов группы может (но не должна обязательно) соответствовать некоммутативность умножения мат-

риц. Для абелевых групп размерности всех неприводимых представлений должны равняться единице.

Неприводимые представления групп играют основную роль в применении теории групп к физическим проблемам. Как будет пояснено ниже, число неприводимых представлений группы равно числу ее классов.

3. Для неприводимых, неэквивалентных, унитарных представлений имеет место следующая *теорема ортогональности* [1]):

$$\sum_R \Gamma_i^*(R)_{\mu\nu}\Gamma_j(R)_{\alpha\beta} = \frac{h}{l_i}\delta_{ij}\delta_{\mu\alpha}\delta_{\nu\beta}. \tag{6.19}$$

Здесь суммирование ведется по всем h элементам группы ($R = A_1, A_2, A_3, \ldots, A_h$); i и j — номера неприводимых представлений, l_i — размерность i-го представления (так как правая часть (6.19) отлична от нуля только при $i = j$, то вместо l_i можно писать l_j); $\Gamma_i^*(R)_{\mu\nu}$ — комплексно-сопряженный $\mu\nu$-й элемент матрицы i-го неприводимого представления для элемента группы R.

Правая часть (6.19) отлична от нуля только при $i = j$, $\mu = \alpha$ и $\nu = \beta$, в этом случае

$$\sum_R |\Gamma_i(R)_{\alpha\beta}|^2 = \frac{h}{l_i}. \tag{6.20}$$

Мы видим, что левая часть от α и β не зависит. Из (6.19) следует, что

$$\sum_R \Gamma_i^*(R)_{\mu\nu}\Gamma_j(R)_{\mu\nu} = 0 \quad \text{при} \quad i \neq j,$$

и

$$\sum_R \Gamma_i^*(R)_{\mu\nu}\Gamma_i(R)_{\alpha\beta} = 0 \quad \text{при} \quad \mu \neq \alpha \text{ или } \nu \neq \beta. \tag{6.21}$$

Из (6.21) видно, что матричные элементы $\Gamma_i(A_1)_{\mu\nu}$, $\Gamma_i(A_2)_{\mu\nu}$, ..., $\Gamma_i(A_h)_{\mu\nu}$ для всех h элементов группы можно рассматривать как компоненты h-мерного вектора, который ортогонален к любому из векторов, полученному для других индексов μ и ν, а также ортогонален к любому из подобных векторов для другого j-го неприводимого представления.

Аналогично в обычном трехмерном пространстве ортогональность трехмерных и двумерных векторов записывается в виде

$$\sum_{i=1}^{3} a_i b_i = \boldsymbol{ab} = 0, \qquad i = x, y, z,$$

$$\sum_{i=1}^{2} a_i b_i = \boldsymbol{ab} = 0, \qquad i = x, y.$$

[1]) Петрашень М. И., Трифонов Е. Д. Применение теории групп в квантовой механике.—М.: Наука, 1967, гл. III, п. 8.

Если имеется всего s неприводимых представлений, то полное число таких взаимно ортогональных векторов равно $\sum_{i=1}^{s} l_i^2$, так как для i-го представления число матричных компонент равно l_i^2. Но в пространстве h измерений нельзя построить больше чем h взаимно ортогональных векторов, поэтому $\sum_{i=1}^{s} l_i^2 \leqslant h$. Можно показать (на чем мы останавливаться не будем), что в этом соотношении выполняется предельное равенство[1]), т. е.

$$\sum_{i=1}^{s} l_i^2 = h. \tag{6.22}$$

Введем важное понятие о шпуре (т. е. сумме диагональных элементов) матрицы представления; в этом случае шпур называется *характером* и обозначается через $\chi_i(R)$. Имеем

$$\chi_i(R) = \operatorname{Sp} \Gamma_i(R) = \sum_{\mu=1}^{l_i} \Gamma_i(R)_{\mu\mu}. \tag{6.23}$$

Здесь l_i — размерность i-го представления. Для характера единичного элемента из (6.23) следует

$$\chi_i(E) = \sum_{\mu=1}^{l_i} \Gamma_i(E)_{\mu\mu} = \sum_{\mu=1}^{l_i} 1 = l_i, \tag{6.24}$$

т. е. характер неприводимого представления единичного элемента группы равен размерности представления.

Из (6.11) следует, что характер (шпур представления) не меняется при преобразовании подобия. Поэтому характеры эквивалентных представлений совпадают. Так как представления элементов одного класса группы связаны преобразованием подобия (см. (2.3)), то $\Gamma(B) = \Gamma(V)^{-1}\Gamma(A)\Gamma(V)$, и характеры для всех элементов класса одинаковы; поэтому можно написать

$$\chi_i(R) = \chi_i(\mathbb{C}_k), \text{ если } R \in \mathbb{C}_k, \tag{6.25}$$

где $\mathbb{C}_k$ символ k-го класса.

Для приложений имеет важное значение следующая теорема *ортогональности характеров:*

$$\sum_R \chi_i^*(R)\,\chi_j(R) = h\delta_{ij}, \tag{6.26}$$

или, учитывая (6.25):

$$\sum_k N_k \chi_i^*(\mathbb{C}_k)\,\chi_j(\mathbb{C}_k) = h\delta_{ij}, \tag{6.27}$$

---

[1]) Хамермеш М. Теория групп.—М.: Мир, 1966, с. 133.

где N_k — число элементов в классе $\mathbb{C}_k$ и суммирование ведется по всем k классам. Для доказательства (6.26) положим в (6.19) $\mu=\nu$ и $\alpha=\beta$, тогда $\delta_{\mu\alpha}\delta_{\nu\beta}=\delta_{\mu\alpha}\delta_{\mu\alpha}=\delta_{\mu\alpha}$ и

$$\sum_R \Gamma_i^*(R)_{\mu\mu}\Gamma_j(R)_{\alpha\alpha}=\frac{h}{l_i}\delta_{ij}\delta_{\mu\alpha}.$$

Суммируя обе части равенства по μ и α, получим

$$\sum_R\sum_\mu \Gamma_i^*(R)_{\mu\mu}\sum_\alpha \Gamma_j(R)_{\alpha\alpha}=\frac{h}{l_i}\delta_{ij}\sum_{\mu\alpha}\delta_{\mu\alpha}.$$

Используя определение характеров (6.23) и учитывая, что $\sum_{\mu\alpha}\delta_{\mu\alpha}=\sum_\mu 1=l_i$ (или $\sum_{\mu\alpha}\delta_{\mu\alpha}=l_j$, что безразлично), и сокращая l_i в правой части, получим (6.26). Из (6.27) следует, что

$$\sum_k N_k\chi_i^*(\mathbb{C}_k)\chi_j(\mathbb{C}_k)=\begin{cases}0, & \text{если } i\neq j,\\ h, & \text{если } i=j.\end{cases} \tag{6.28}$$

Заметим, что это выражение может служить критерием неприводимости представления. Представление Γ_i неприводимо тогда и только тогда, когда

$$\sum_k N_k|\chi_i(\mathbb{C}_k)|^2=h. \tag{6.29}$$

Если представление $\Gamma(R)$ содержит, например, два неприводимых представления I и II, то очевидно, что для него $\chi(R)=\chi^{\mathrm{I}}(R)+\chi^{\mathrm{II}}(R)$ (см. ниже (6.32)). Поэтому для него сумма $\sum_R|\chi(R)|^2$ равна $2h$, если представления I и II неэквивалентны и их характеры ортогональны, и равна $4h$, если они эквивалентны. То есть условие (6.29) является достаточным и необходимым условием неприводимости.

Будем теперь рассматривать отнормированные характеры классов $\sqrt{N_1}\chi_i(\mathbb{C}_1)$, $\sqrt{N_2}\chi_i(\mathbb{C}_2)$, ..., $\sqrt{N_k}\chi_i(\mathbb{C}_k)$ как составляющие вектора в пространстве классов (аналогично тому, как это делалось для составляющих $\Gamma_i(R)_{\mu\nu}$). Тогда из условия взаимной ортогональности таких векторов (6.27) следует, что число неприводимых представлений меньше или равно числу классов. И в этом случае можно доказать[1]), что выполняется предельное соотношение, т. е.

$$\text{число неприводимых представлений}=\text{числу классов} \tag{6.30}$$

Можно вывести[2]) соотношение, аналогичное (6.27), имеющее

[1]) Tinkham M. Group Theory and Quantum Mechanics.—McGraw-Hill Book Co, 1964, гл. 3.

[2]) См. там же.

вид

$$\sum_i \chi_i^* (\mathbb{C}_k)\, \chi_i (\mathbb{C}_l) = \frac{h}{N_k}\, \delta_{kl}, \tag{6.31}$$

где ортогональность характеров имеет место не между разными представлениями, а между разными классами.

Хотя характеры неприводимых представлений дают меньше информации, чем сами неприводимые представления, во многих случаях они достаточны для решения физических задач. Поэтому представляется заманчивым получить характеры неприводимых представлений разных классов, не определяя в явном виде сами неприводимые представления. Это можно сделать, воспользовавшись следующими соотношениями: (6.30), (6.22), (6.27), (6.31).

Таблицу характеров мы будем строить так, чтобы каждой строке соответствовало определенное неприводимое представление, а каждому столбцу определенный класс. Из (6.30) следует, что число строк и столбцов должно быть одинаковым, т. е. таблица характеров квадратная. Если первой строке соответствует единичное представление, то характеры ее для всех классов равны единице. Если первому столбцу соответствует класс E, то характеры неприводимых представлений, согласно (6.24), равны их размерностям l_i, которые могут быть определены из (6.22) (обычно при заданном h и s величины l_i определяются однозначно). Зная первую строку и первый столбец таблицы, можно остальные характеры определять подбором так, чтобы удовлетворялись условия ортонормированности характеров по строкам (6.27) и по столбцам (6.31). Обычно этих условий хватает с избытком, так что часть их служит для проверки таблицы [1]).

Используя сформулированные выше свойства таблицы характеров, легко построить ее для любой группы второго порядка. В этом случае группа абелева и оба неприводимых представления, как это следует из (6.22), имеют размерность единица. Используя условие ортогональности строк (или столбцов), получим для групп $\boldsymbol{C}_i = \{E, J\}$ и $\boldsymbol{C}_s = \boldsymbol{C}_{1h} = \{E, \sigma_h\}$ табл. II.3.

Проиллюстрируем еще приведенные выше соотношения на примере группы $\boldsymbol{D}_3$ (§ 2, п. 2). Эта неабелева группа шестого порядка ($h = 6$) состоит из трех классов:

$$\mathbb{C}_1 = E, \quad \mathbb{C}_2 = \{A, B, C\} = 3C_2, \quad \mathbb{C}_3 = \{D, F\} = 2\,(C_3; C_3^2).$$

Так как число неприводимых представлений этой группы тоже должно равняться трем, то три неэквивалентные представления (6.16) являются неприводимыми представлениями. Мы видим, что

---

[1]) Не существует определенного алгоритма для составления таблицы характеров, так что в этом деле нужна некоторая сноровка; правда, нужно сказать, что для интересующих физиков групп такие таблицы имеются.

Таблица II.3

C_i	E	J
C_s	E	σ_h
Γ_1	1	1
Γ_2	1	—1

Таблица II.4

D_3	E	$3C_2$	$2(C_3; C_3^2)$
Γ_1	1	1	1
Γ_2	1	—1	1
Γ_3	2	0	—1

для каждого представления характеры, относящиеся к одному классу, одинаковы. В табл. II.4 даны характеры группы $\boldsymbol{D}_3$ (множители при $\mathbb{C}_2$ и $\mathbb{C}_3$, напоминают о числе элементов в классе). Из таблицы видно, что сумма квадратов размерностей представлений равна порядку группы (6.22): $1^2+1^2+2^2=6$. Легко проверить выполнимость (6.27), например, условие нормировки для Γ_3 имеет вид: $2^2+3\cdot 0+2\cdot(-1)^2=6$; условие ортогональности между Γ_2 и Γ_3: $1\cdot 2+3\cdot(-1)\cdot 0+2\cdot 1\cdot(-1)=0$. Так же можно проверить выполнимость (6.31); например, нормировка столбца $\mathbb{C}_2$: $1^2+(-1)^2+0=6/3=2$, а ортогональность между $\mathbb{C}_1$ и $\mathbb{C}_3$: $1\times 1+1\cdot 1+2\cdot(-1)=0$. Можно проверить и остальные соотношения.

4. Представим себе некоторое приводимое представление Γ. Посредством соответствующего преобразования подобия все матрицы приводимого представления $\Gamma(R)$ (R—элемент группы) могут быть одновременно приведены к квазидиагональным матрицам одинаковой структуры, в которых неприводимое представление $\Gamma_i(R)$ встречается a_i раз[1]). Мы будем писать

$$\Gamma=a_1\Gamma_1+a_2\Gamma_2+\ldots+a_s\Gamma_s=\sum_i a_i\Gamma_i, \qquad (6.32)$$

понимая условность этой записи (справа у нас стоит «сумма» матриц разного ранга; она иногда называется *прямой* суммой). Так как при таком преобразовании подобия характер матрицы не меняется, то

$$\chi(R)=\sum_i a_i\chi_i(R) \qquad (6.33)$$

(здесь уже сумма и равенство не носят условного характера). Пользуясь условием ортогональности характеров неприводимых представлений (6.26), получим из (6.33)

$$\sum_R \chi(R)\,\chi_j^*(R)=\sum_i a_i\sum_R \chi_i(R)\,\chi_j^*(R)=\sum_i a_i h\delta_{ij}=ha_j,$$

[1]) Можно доказать однозначность такого разложения; конечно, некоторые a_i могут оказаться равными единице или нулю.

откуда

$$a_j = \frac{1}{h}\sum_R \chi(R)\,\chi_j^*(R) = \frac{1}{h}\sum_{k=1}^{s} N_k \chi(\mathbb{C}_k)\,\chi_j^*(\mathbb{C}_k), \qquad (6.34)$$

где N_k — число элементов в классе $\mathbb{C}_k$.

Эта формула, определяющая, сколько раз данное неприводимое представление Γ_j встречается в приводимом представлении Γ, играет важную роль во всех приложениях.

5. В § 2 п. 5 мы ввели понятие о прямом произведении групп

$$\mathfrak{A} = \{A_1 \equiv E, A_2, A_3, \ldots, A_{h_a}\} \text{ и } \mathfrak{B} = \{B_1 \equiv E, B_2, B_3, \ldots, B_{h_3}\},$$

которое мы обозначили как $\mathfrak{A}\times\mathfrak{B}$.

Пусть Γ_1 — некоторое представление группы $\mathfrak{A}$, а Γ_2 — представление группы $\mathfrak{B}$, вообще говоря, другой размерности. Покажем, что прямое произведение (см. Приложение 3, п. 6) представлений $\Gamma_1(A_k)\times\Gamma_2(B_l)$ есть представление прямого произведения групп $\mathfrak{A}\times\mathfrak{B}$. Пусть элементы группы прямого произведения перемножаются по закону

$$(A_kB_l)(A_{k'}B_{l'}) = (A_kA_{k'})(B_lB_{l'}) = A_{k''}B_{l''}, \qquad (6.35)$$

где

$$A_{k''} = A_kA_{k'} \text{ и } B_{l''} = B_lB_{l'}.$$

Покажем, что этому же закону умножения удовлетворяет прямое произведение представлений $\Gamma_1\times\Gamma_2$; имеем

$$\begin{aligned}[\Gamma_1(A_k)\times\Gamma_2(B_l)][\Gamma_1(A_{k'})\times\Gamma_2(B_{l'})] &= \\ = \Gamma_1(A_k)\,\Gamma_1(A_{k'})\times\Gamma_2(B_l)\,\Gamma_2(B_{l'}) &= \\ = \Gamma_1(A_kA_{k'})\times\Gamma_2(B_lB_{l'}) = \Gamma_1(A_{k''})\times\Gamma_2(B_{l''}). & \qquad (6.36)\end{aligned}$$

Здесь первое равенство следует из (П.3.50), а последнее из (6.35). Мы видим, что прямые произведения $\Gamma_1\times\Gamma_2$ удовлетворяют закону умножения (6.35).

Из (6.36) следует, что представление прямого произведения групп равно прямому произведению представлений, т. е.

$$\Gamma(A_kB_l) = \Gamma_1(A_k)\times\Gamma_2(B_l). \qquad (6.37)$$

Используя (П.3.52), получим

$$\mathrm{Sp}\,\Gamma(A_k, B_l) = \mathrm{Sp}\,[\Gamma_1(A_k)\times\Gamma_2(B_l)] = \mathrm{Sp}\,\Gamma_1(A_k)\cdot\mathrm{Sp}\,\Gamma_2(B_l),$$

или

$$\chi(A_kB_l) = \chi_1(A_k)\,\chi_2(B_l), \qquad (6.38)$$

т. е. характер представления прямого произведения групп для элемента A_kB_l равен произведению характеров представлений для элементов A_k и B_l.

То, что это соотношение связывает характеры неприводимых представлений, может быть обосновано следующим образом. Из

(6.38) и (6.26) следует, что

$$\sum_{k,\,l} |\chi(A_k B_l)|^2 = \sum_k |\chi_1(A_k)|^2 \sum_l |\chi_2(B_l)|^2 = h_a h_b = h,$$

где h_a, h_b и h число элементов групп $\mathfrak{A}$, $\mathfrak{B}$ и их прямого произведения $\mathfrak{A}\times\mathfrak{B}$. Таким образом, если $\mathfrak{A}$ и $\mathfrak{B}$ — неприводимые представления групп $\mathfrak{A}$ и $\mathfrak{B}$, то $\Gamma_1\times\Gamma_2$ тоже неприводимое представление группы $\mathfrak{A}\times\mathfrak{B}$ (так как только для неприводимого представления выполняется соотношение (6.26).

Соотношение (6.38) облегчает составление таблиц характеров для тех групп, которые можно рассматривать как прямые произведения более простых групп.

В § 3, п. 4 мы видели, что группа $\boldsymbol{D}_{nh} = \boldsymbol{D}_n \times \boldsymbol{C}_s$, где группа $\boldsymbol{C}_s = \{E_1\sigma_h\}$. В табл. II.4 приведены характеры группы $\boldsymbol{D}_3$. В группе $\boldsymbol{D}_{3h}$ вдвое большее число элементов, классов и представлений, чем в группе $\boldsymbol{D}_3$. Используя формулу (6.38), легко построить таблицу характеров группы $\boldsymbol{D}_{3h}$ (табл. II.5). Вообще, если рассматриваемая группа $\boldsymbol{G}$ есть прямое произведение некоторой группы $\boldsymbol{\mathcal{H}}$ на группу второго порядка $\boldsymbol{C}_s$ или $\boldsymbol{C}_i$ с характерами табл. II.3, то таблица характеров группы $\boldsymbol{G}$ определяется по схеме табл. II.6, где χ — характеры группы $\boldsymbol{\mathcal{H}}$.

Таблица II.5

D_{3h}	E	$3C_2$	$2(C_3; C_3^2)$	σ_h	$3\sigma_h C_2$	$2\sigma_h(C_3; C_3^2)$
Γ_{1+}	1	1	1	1	1	1
Γ_{2+}	1	—1	1	1	—1	1
Γ_{3+}	2	0	—1	2	0	—1
Γ_{1-}	1	1	1	—1	—1	—1
Γ_{2-}	1	—1	1	—1	1	—1
Γ_{3-}	2	0	—1	—2	0	1

В табл. II.7 приведены характеры группы $\boldsymbol{O}$ и изоморфной ей группы $\boldsymbol{T}_d$. Обозначения неприводимых представлений в первом и втором столбцах даны соответственно по Баукерту,

Таблица II.6

G	Классы	Классы
Неприводимые представления	χ	χ
Неприводимые представления	χ	$-\chi$

Таблица II.7

O		T_d		E	$8C_3$	C_{24}^2	$6C_2$	$6C_4$
				E	$8C_3$	$3C_4^2$	6σ	$6S_4$
Γ_1	A_1	Γ_1	A_1	1	1	1	1	1
Γ_2	A_2	Γ_2	A_2	1	1	1	-1	-1
Γ_{12}	E	Γ_{12}	E	2	-1	2	0	0
Γ_{15}	F_1 xyz	Γ_{25}	F_1	3	0	-1	-1	1
Γ_{25}	F_2	Γ_{15}	F_2 xyz	3	0	-1	1	-1

Смолуховскому и Вигнеру и по статьям, посвященным молекулярным спектрам (в последнем случае принято обозначать через A одномерные, E — двумерные и F (или T) — трехмерные представления).

Поскольку полная группа симметрии куба $\boldsymbol{O}_h = \boldsymbol{O} \times \boldsymbol{C}_i$, где $\boldsymbol{C}_i = \{E, J\}$, таблица характеров группы $\boldsymbol{O}_h$ может быть получена из табл. II.7 посредством применения схемы табл. II.6.

Если для некоторой точечной группы элементы R и R^{-1} относятся к разным классам (это имеет, например, место, если ось вращения для элемента R не является двусторонней), тогда для некоторого неприводимого представления

$$\Gamma(R^{-1}) = \Gamma^{-1}(R) = \Gamma^{+}(R) = \tilde{\Gamma}^{*}(R),$$

так как Γ — унитарная матрица; отсюда

$$\Gamma(R^{-1})_{ik} = \Gamma^*(R)_{ki},$$

и, следовательно,

$$\chi(R^{-1}) = \sum_i \Gamma(R^{-1})_{ii} = \sum_i \Gamma^*(R)_{ii} = \chi^*(R),$$

или

$$\chi(R^{-1}) = \chi^*(R), \tag{6.39}$$

т. е. характеры элементов R и R^{-1} комплексно сопряжены.

Таблица II.8

C_3	E	C_3	C_3^2
Γ_1	1	1	1
Γ_2	1	ω	ω^2
Γ_3	1	ω^2	ω

$\omega = e^{2\pi i/3}$

В табл. II.8 даны характеры группы C_3, для которой элементы и C_3 и C_3^{-1} относятся к разным классам [1]). Мы видим, что действительно для представления Γ_2

$$\chi(C_3) = \omega = e^{2\pi i/3} \text{ и } \chi(C_3^{-1}) = \omega^2 = e^{4\pi i/3} = e^{-2\pi i/3},$$

т. е.

$$\chi(C_3) = \chi^*(C_3^{-1}).$$

Аналогичная ситуация имеет место для неприводимого представления Γ_3.

§ 7. Квантовая механика и теория групп

1. В начале § 6 мы уже указали, что матричные представления групп симметрии естественно возникают при рассмотрении физических проблем.

Важнейшим примером этого является рассмотрение решений стационарного уравнения Шредингера.

Представим себе физическую систему в конфигурационном пространстве n координат $x_1, x_2, \ldots, x_n$, образующих n-мерный

[1]) Более полные таблицы характеров точечных групп, см., например, у Ландау Л. Д., Лифшиц Е. М. Квантовая механика.—М.: Наука, 1974, с. 434.

вектор $\boldsymbol{x}=\{x_1, x_2, \ldots, x_n\}$. Волновая функция системы $\psi(\boldsymbol{x})$ в стационарном состоянии удовлетворяет уравнению Шредингера

$$\hat{\mathcal{H}}(\boldsymbol{x})\,\psi(\boldsymbol{x})=\mathcal{E}\,\psi(\boldsymbol{x}), \tag{7.1}$$

где $\hat{\mathcal{H}}(\boldsymbol{x})$ — гамильтониан системы, $\mathcal{E}$ — собственное значение ее энергии.

Если состояние с энергией $\mathcal{E}$ l-кратно вырождено, то ему соответствуют l линейно независимых собственных функций:

$$\psi_1(\boldsymbol{x}),\ \psi_2(\boldsymbol{x}),\ \ldots,\ \psi_i(\boldsymbol{x}),\ \ldots,\ \psi_l(\boldsymbol{x}). \tag{7.2}$$

Из квантовой механики известно, что эти функции всегда можно считать ортонормированными [1]), т. е. скалярное произведение

$$(\psi_i,\ \psi_k)=\int \psi_i^*(\boldsymbol{x})\,\psi_k(\boldsymbol{x})\,d\boldsymbol{x}=\delta_{ik}, \tag{7.3}$$

где $d\boldsymbol{x}=dx_1\,dx_2\,\ldots\,dx_n$.

Любое решение уравнения (7.1), соответствующее собственному значению $\mathcal{E}$, может быть представлено как линейная комбинация функций (7.2).

Представим себе, что рассматриваемая система обладает некоторой симметрией, например пространственной симметрией некоторой точечной группы или симметрией по отношению к перестановкам тождественных частиц между собой.

Всякая операция такой симметрии (вращение, отражение, инверсия, перестановка частиц) связана с некоторым линейным преобразованием конфигурационных координат системы, которое может быть записано в виде

$$\boldsymbol{x}'=\boldsymbol{R}\boldsymbol{x}, \tag{7.4}$$

где $\boldsymbol{R}$ — вещественная ортогональная матрица линейного преобразования. Обратное (7.4) преобразование имеет вид

$$\boldsymbol{x}=\boldsymbol{R}^{-1}\boldsymbol{x}', \tag{7.5}$$

где $\boldsymbol{R}^{-1}$ — обратная матрица. Для вещественных ортогональных матриц $\boldsymbol{R}^{-1}=\tilde{\boldsymbol{R}}$, т. е. $(\boldsymbol{R}^{-1})_{ij}=\boldsymbol{R}_{ji}$ (см. (П.3.28)). Например, равносторонний треугольник (рис. II.1) может быть описан шестью координатами x_1, x_2, x_3, x_4, x_5 и x_6, которые последовательно равны координатам x и y вершин *1*, *2* и *3*.

Операции D (вращению на 120° по часовой стрелке вокруг оси, перпендикулярной к плоскости треугольника и проходящей

---

[1]) См. в Приложении 3 п. 5 аналогичный вопрос об ортогонализации собственных векторов, соответствующих одному собственному значению эрмитовых матриц.

через его центр) соответствует преобразование [1])

$$\left.\begin{aligned}
x_1' &= x_3 = 0\cdot x_1 + 0\cdot x_2 + 1\cdot x_3 + 0\cdot x_4 + 0\cdot x_5 + 0\cdot x_6,\\
x_2' &= x_4 = 0\cdot x_1 + 0\cdot x_2 + 0\cdot x_3 + 1\cdot x_4 + 0\cdot x_5 + 0\cdot x_6,\\
x_3' &= x_5 = \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 1\cdot x_5 + \ldots,\\
x_4' &= x_6 = \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots + 1\cdot x_6,\\
x_5' &= x_1 = 1\cdot x_1 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots .,\\
x_6' &= x_2 = \ldots + 1\cdot x_2 + \ldots\ldots\ldots\ldots\ldots .,
\end{aligned}\right\} \tag{7.6}$$

так, что матрица ортогонального преобразования

$$\boldsymbol{R} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{7.7}$$

Легко убедиться в том, что обратное преобразование

$$\boldsymbol{R}^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{7.8}$$

и что, действительно,

$$\boldsymbol{R}^{-1} = \tilde{\boldsymbol{R}}. \tag{7.9}$$

Если система обладает некоторой симметрией, так что она совпадает сама с собой при преобразовании координат (7.5), то потенциальная энергия системы $V(\boldsymbol{x})$, входящая в гамильтониан $\hat{\mathcal{H}}(\boldsymbol{x})$ в (7.1), удовлетворяет условию

$$V(\boldsymbol{x}) = V(\boldsymbol{R}^{-1}\boldsymbol{x}') = V(\boldsymbol{x}'), \tag{7.10}$$

где мы в $V(\boldsymbol{x})$ произвели преобразование (7.5) от $\boldsymbol{x}$ к $\boldsymbol{x}'$. Например, потенциальная энергия электрона в поле ядра атома Ze^2/r обладает сферической симметрией, т. е. не меняется при любых поворотах атома (или координатной системы) вокруг начала координат, совпадающего с ядром. Действительно, производя ортогональное преобразование (7.5) по отношению к координатам электрона $x_1 = x$, $x_2 = y$ и $x_3 = z$, получим

$$\frac{Ze^2}{\sqrt{x^2+y^2+z^2}} = \frac{Ze^2}{\sqrt{x'^2+y'^2+z'^2}}. \tag{7.11}$$

[1]) При этом вершина треугольника *1* переходит в вершину *2*, *2* в *3* и *3* в *1*.

Так как оператор Лапласа инвариантен относительно ортогональных преобразований (7.5), то

$$\hat{\mathscr{H}}(\boldsymbol{x}) = \hat{\mathscr{H}}(\boldsymbol{R}^{-1}\boldsymbol{x}') = \hat{\mathscr{H}}(\boldsymbol{x}'). \tag{7.12}$$

Если мы произведем замену координат (7.5) во всем уравнении (7.1), то получим

$$\hat{\mathscr{H}}(\boldsymbol{x}')\,\psi(\boldsymbol{R}^{-1}\boldsymbol{x}') = \mathscr{E}\,\psi(\boldsymbol{R}^{-1}\boldsymbol{x}'),$$

где мы переменные $\boldsymbol{x}'$ можем, конечно, обозначить через $\boldsymbol{x}$:

$$\hat{\mathscr{H}}(\boldsymbol{x})\,\psi(\boldsymbol{R}^{-1}\boldsymbol{x}) = \mathscr{E}\psi(\boldsymbol{R}^{-1}\boldsymbol{x}). \tag{7.13}$$

Все преобразования симметрии, оставляющие гамильтониан инвариантным, носят название *группы уравнения Шредингера*.

Сравнивая (7.13) с (7.1), мы видим, что функция $\psi(\boldsymbol{R}^{-1}\boldsymbol{x}) \equiv \equiv \varphi(\boldsymbol{x})$ удовлетворяет тому же уравнению, что и функция $\psi(\boldsymbol{x})$ при том же собственном значении энергии $\mathscr{E}$.

Если собственное значение $\mathscr{E}$ не вырождено, то $\varphi(\boldsymbol{x})$ может отличаться от $\psi(\boldsymbol{x})$ только постоянным множителем, т. е.

$$\varphi(\boldsymbol{x}) = c\psi(\boldsymbol{x}). \tag{7.14}$$

Если $R^2 = E$, то $c^2 = 1$ и, следовательно, $c = \pm 1$. Таким образом, в случае $R^2 = E$ существуют два решения уравнения (7.1): одно, не меняющее свой знак при преобразовании симметрии (симметричное), и другое, меняющее свой знак на обратный при преобразовании симметрии (антисимметричное) [1]).

Если же собственное значение $\mathscr{E}$ l-кратно вырождено, так что $\psi(\boldsymbol{x}) = \psi_i(\boldsymbol{x})$ и равна одной из функций (7.2), то $\psi_i(\boldsymbol{R}^{-1}\boldsymbol{x})$ (см. (7.13)) является решением уравнения Шредингера (7.1) для того же собственного значения энергии $\mathscr{E}$, поэтому оно должно выражаться линейно через волновые функции (7.2):

$$\psi_i(\boldsymbol{R}^{-1}\boldsymbol{x}) = \hat{P}_R\psi_i(\boldsymbol{x}) = \sum_{k=1}^{l} \Gamma(R)_{ki}\,\psi_k(\boldsymbol{x}). \tag{7.15}$$

Это выражение совпадает с (6.2) для преобразования базисных функций ψ_i. Из (6.3) следует, что элементы матрицы $\Gamma(R)_{ki}$ — матричные элементы оператора $\hat{P}_R$, построенные на волновых (базисных) функциях (7.2). Из (6.5а) следует, что унитарные матрицы $\Gamma(R)$ в (7.15) осуществляют представление группы уравнения Шредингера (7.1).

Мы можем сделать важное заключение, что каждому собственному значению энергии $\mathscr{E}$ соответствует определенное (с точностью до преобразования подобия) представление группы уравнения Шредингера.

---

[1]) См. решение для атома гелия во введении (§ 1).

В дальнейшем мы будем считать, что если нет *случайного вырождения*, то представление группы уравнения Шредингера, соответствующее определенному собственному значению энергии, является неприводимым.

В самом деле, если бы это представление было приводимым, то матрицы $\Gamma(R)$ можно было бы соответствующим преобразованием подобия (или, иначе, соответствующим выбором l базисных функций (7.2)) привести к квазидиагональному виду, например к двум блокам рангов n и $l-n$, расположенных вдоль главной диагонали матрицы $\Gamma(R)$. В этом случае n функций из (7.2) преобразовывались бы, при всех преобразованиях симметрии R, только друг через друга и то же имело бы место для остальных $l-n$ функций. Таким образом, каждая из двух групп функций вела бы себя так, как будто она принадлежала некоторому определенному уровню энергии, но тогда было бы маловероятным, и, во всяком случае, не связанным с симметрией гамильтониана $\hat{\mathcal{H}}(\boldsymbol{x})$ такое случайное совпадение обоих уровней энергии (*случайное вырождение*).

Таким образом, мы будем считать, что каждому уровню энергии соответствует определенное неприводимое представление группы уравнения Шредингера. Конечно, разным уровням энергии может соответствовать одно и то же неприводимое представление. Например, мы увидим дальше, что неприводимые представления электрона в атоме, в одночастичном приближении, характеризуются орбитальным квантовым числом l, так что s-состояние ($l=0$), p-состояние ($l=1$), d-состояние ($l=2$) и т. д. — различные неприводимые представления непрерывной группы вращений. При этом каждому уровню энергии электрона соответствует свое неприводимое представление с заданным l. Но, конечно, разным уровням энергии с различными главными квантовыми числами n могут соответствовать одни и те же неприводимые представления с одинаковым l.

Рассмотрим атом в узле кубической кристаллической решетки, и пусть результирующее поле, действующее на атом, обладает симметрией группы $\boldsymbol{O}$ (приближение, когда действие всех остальных электронов и ядер решетки можно заменить электрическим полем, обладающим симметрией решетки), которая имеет пять неприводимых представлений (табл. II.7). Тогда все уровни электрона атома в узле кубического кристалла должны принадлежать к одному из пяти неприводимых представлений группы $\boldsymbol{O}$. Таким образом, можно сказать, что в поле симметрии $\boldsymbol{O}$ не могут существовать уровни энергии с кратностью вырождения, большей трех.

В следующем параграфе мы рассмотрим вопрос, как расщепляются уровни энергии валентного атома при действии на него

кристаллического поля, обладающего симметрией $\boldsymbol{O}$ или $\boldsymbol{O}_h$ [1]).

2. Рассмотрим важный случай, когда гамильтониан $\hat{\mathscr{H}}$ в (7.1) может быть представлен как сумма

$$\hat{\mathscr{H}} = \hat{\mathscr{H}}_0 + \hat{\mathscr{H}}', \tag{7.16}$$

причем симметрия $\hat{\mathscr{H}}'$ ниже симметрии $\hat{\mathscr{H}}_0$, так что группа симметрии $\hat{\mathscr{H}}'$ является подгруппой группы симметрии $\hat{\mathscr{H}}_0$. Такая ситуация возникает, например, если наложить внешнее однородное электрическое поле E_0 на атом, в результате чего сферическая симметрия для электронов атома нарушается возмущением $\hat{\mathscr{H}}' = -eE_0 z$ ($\boldsymbol{E}_0 \parallel$ оси z).

Аналогичное понижение симметрии имеет место, если поместить атом в один из узлов кристаллической решетки и считать, что на электрон в атоме, помимо центральных сил, действующих со стороны ядра атома и его остальных электронов, дополнительно действует усредненное электрическое поле, определяемое симметрией решетки. Очевидно, что симметрия полного гамильтониана определяется его менее симметричной частью $\hat{\mathscr{H}}'$.

Рассмотрим, что произойдет с некоторым, вообще говоря, вырожденным уровнем энергии $\mathscr{E}_0$, соответствующим гамильтониану $\hat{\mathscr{H}}_0$, при «включении» гамильтониана $\hat{\mathscr{H}}'$. Поскольку группа $\mathscr{H}'$ [2]) является подгруппой группы $\hat{\mathscr{H}}_0$, неприводимое представление уровня $\mathscr{E}_0$ будет являться представлением, вообще говоря, приводимым группы $\hat{\mathscr{H}}'$. Это приводимое представление мы можем разложить на неприводимые представления группы $\hat{\mathscr{H}}'$, и тогда, если энергия $\hat{\mathscr{H}}'$ много меньше расстояния между термами энергии невозмущенного гамильтониана $\hat{\mathscr{H}}_0$, мы можем сказать, на сколько уровней и какой кратности вырождения расщепится уровень $\mathscr{E}_0$ под действием поля $\hat{\mathscr{H}}'$. Однако мы ничего не можем сказать ни о величине этого расщепления, ни о том, в какой последовательности расположены расщепленные уровни.

Пусть невозмущенная система, т. е. гамильтониан $\hat{\mathscr{H}}_0$, обладает симметрией $\boldsymbol{O}$ (см. табл. II.7). Нас интересует, что произойдет с неприводимым представлением F_2 (трехкратно вырожденным уровнем F_2) при наложении поля $\hat{\mathscr{H}}'$ с симметрией $\boldsymbol{D}_3$ (см. табл. II.4), если его ось C_3 совпадает с одной из осей C_3 группы $\boldsymbol{O}$. Можно увидеть, что элементы группы $\boldsymbol{O}$: E, $2(C_3, C_3^2)$, $3C_2$ образуют подгруппу $\boldsymbol{D}_3$. Представление F_2 является приво-

[1]) Как мы покажем дальше, группа $\boldsymbol{O}$ обладает в этом случае, с точки зрения характера расщепления термов, теми же свойствами, что и группа $\boldsymbol{O}_h = \boldsymbol{O} \times C_i$.

[2]) Мы говорим для краткости «группа $\hat{\mathscr{H}}'$» вместо «группа симметрии $\hat{\mathscr{H}}'$».

димым представлением этой подгруппы (это видно уже из того, что подгруппа $\boldsymbol{D}_3$ не имеет неприводимых представлений с размерностью, равной трем). Итак, мы должны приводимое (для $\boldsymbol{D}_3$) представление F_2 с характерами 3, 0, 1 (для элементов E, $8\,(C_3, C_3^2)$ и $6C_2$) разложить на неприводимые представления Γ_1, Γ_2 и Γ_3 группы $\boldsymbol{D}_3$. По формулам (6.32) и (6.34) получим

$$F_2 = a_1\Gamma_1 + a_2\Gamma_2 + a_3\Gamma_3$$

и

$$a_1 = \frac{1}{6}\left[3\cdot 1 + 2\cdot 0\cdot 1 + 3\cdot 1\cdot 1\right] = 1,$$

$$a_2 = \frac{1}{6}\left[3\cdot 1 + 2\cdot 0\cdot 1 + 3\cdot 1\cdot(-1)\right] = 0,$$

$$a_3 = \frac{1}{6}\left[3\cdot 2 + 2\cdot 0\cdot(-1) + 3\cdot 1\cdot 0\right] = 1.$$

Таким образом,

$$F_2 = \Gamma_1 + \Gamma_3,$$

т. е. трехкратно вырожденный уровень F_2, под действием поля симметрии $\boldsymbol{D}_3$ расщепляется на невырожденный уровень Γ_1 и двукратно вырожденный уровень Γ_3.

3. Наряду с прямым произведением двух групп введем важное понятие о прямом произведении неприводимых представлений одной группы.

Пусть имеется два неприводимых представления группы Γ_1 и Γ_2 размерностей l и m. Базисные функции этих неприводимых представлений:

$$\Gamma_1\colon \psi_1, \psi_2, \ldots, \psi_i, \ldots, \psi_l,\ \Gamma_2\colon \varphi_1, \varphi_2, \ldots, \varphi_k, \ldots, \varphi_m. \quad (7.17)$$

Легко видеть, что произведения $\psi_i\varphi_k$ $(i = 1, 2, \ldots, l, k = 1, 2, \ldots, m)$ тоже образуют базис представления группы размерности $l\cdot m$. В самом деле,

$$\hat{P}_R(\psi_i\varphi_k) = \hat{P}_R\psi_i\cdot\hat{P}_R\varphi_k = \sum_j \Gamma_i(R)_{ji}\psi_j \sum_p \Gamma_2(R)_{pk}\varphi_p =$$

$$= \sum_{j,p}\left[\Gamma_1(R)_{ji}\Gamma_2(R)_{pk}\right]\psi_j\varphi_p = \sum_{j,p}\Gamma(R)_{jp,ik}\psi_j\varphi_p,$$

где матрица $\Gamma(R) = \Gamma_1(R)\times\Gamma_2(R)$, т. е. равна прямому произведению матриц Γ_1 и Γ_2 (см. Приложение 3, п. 6).

Отсюда, совершенно так же как в § 6, п. 5, следует, что прямое произведение матриц неприводимых представлений группы есть тоже ее представление. Однако, в то время как прямое произведение неприводимых представлений двух разных групп есть тоже неприводимое представление прямого произведения групп, — прямое произведение неприводимых представлений одной группы есть, вообще говоря, ее приводимое представление.

Аналогично (6.38) можно показать, что характер $\chi(R)$ представления Γ равен произведению характеров $\chi_1(R)$ и $\chi_2(R)$ не-

приводимых представлений Γ_1 и Γ_2, т. е.

$$\chi(R) = \chi_1(R)\,\chi_2(R). \tag{7.18}$$

Приводимое представление $\Gamma = \Gamma_1 \times \Gamma_2$ может быть разложено по формулам (6.32), (6.34) на неприводимые представления группы.

§ 8. Применение теории групп к исследованию расщепления уровней энергии примесного атома в кристалле и к классификации нормальных колебаний многоатомной молекулы

Рассмотрим два примера применения теории групп в физике: 1) расщепление энергетических уровней атома (иона) в поле кристалла и 2) исследование собственных колебаний молекулы.

Первый из этих вопросов имеет прямое отношение к исследованию примесных атомов в твердом теле, второй — представляет собой не только поучительный пример применения теории групп в физике, но и существен для изучения колебаний кристаллической решетки.

1. Для исследования первого вопроса предварительно рассмотрим непрерывную группу вращений.

Наряду с конечными точечными группами, рассмотренными в § 3, существуют так называемые непрерывные точечные группы с бесконечным числом элементов. Это группы *аксиальной* и *сферической* симметрии. Мы остановимся только на группе сферической симметрии $\mathcal{R}_i$, которая состоит из бесконечного числа поворотов на произвольные углы вокруг любых осей, проходящих через неподвижную точку O, и отражений в любой плоскости, содержащей ту же точку O. Так как $\sigma_h C_2 = J$-инверсии, то можно сказать, что группа сферической симметрии состоит из произвольных вращений вокруг всех осей, проходящих через центр, и инверсии. Заметим, что подгруппа сферической группы, состоящая только из всех произвольных вращений, получила название группы *собственных вращений* $\mathcal{R}$ или просто группы вращений (см. (П.3.33) и дальше), в отличие от сферической группы (включающей инверсию), которую называют также группой *несобственных вращений*. Очевидно, что $\mathcal{R}_i = \mathcal{R} \times \boldsymbol{C}_i$, где $\boldsymbol{C}_i = \{E, J\}$.

Очевидно, что в группе вращений $\mathcal{R}$ все оси эквивалентны и двусторонни; поэтому классами этой группы являются повороты на заданный по абсолютной величине угол α вокруг любой оси. Классы группы $\mathcal{R}_i$ получаются непосредственно из прямого произведения $\mathcal{R} \times \boldsymbol{C}_i$.

Для определения характеров непрерывной группы вращений $\mathcal{R}$, рассмотрим решения уравнения Шредингера с гамильтонианом, инвариантным относительно группы $\mathcal{R}$. Рассмотрим электрон

в поле сферической симметрии, для которого гамильтониан

$$\hat{\mathscr{H}}(\boldsymbol{x}) = -\frac{\hbar^2}{2m}\nabla^2 + V(r), \tag{8.1}$$

где потенциальная энергия электрона $V(r) = V\left(\sqrt{x^2+y^2+z^2}\right)$ обладает сферической симметрией [1]).

Волновые функции уравнения Шредингера с гамильтонианом (8.1) имеют вид [2])

$$\psi_{nlm}(r,\ \vartheta,\ \varphi) = R_{nl}(r)\, Y_{lm}(\vartheta,\ \varphi). \tag{8.2}$$

Здесь r, ϑ, φ — полярные координаты электрона, n, l, m — главное, азимутальное и магнитное квантовые числа, $R_{nl}(r)$ — радиальная часть волновой функции, зависящая от конкретного вида потенциала $V(r)$ и

$$Y_{lm}(\vartheta,\ \varphi) = P_{lm}(\cos\vartheta)\, e^{im\varphi} \tag{8.3}$$

— сферическая функция, где $P_{lm}(\cos\vartheta)$ — полином степени l от $\cos\vartheta$, непосредственно выражающийся через полином Лежандра $P_l(\cos\vartheta)$. Собственные значения энергии $\mathscr{E}_{nl}$ $(2l+1)$-кратно вырождены по магнитному квантовому числу m $(m = -l, -l+1, \ldots, 1, \ldots, l)$.

Для определения неприводимого представления, соответствующего повороту вокруг любой оси на угол α, выберем в качестве оси поворота ось z, тогда в соответствии с (7.15) получим

$$\hat{P}_\alpha \psi_{nlm} = \psi_{nlm}(\alpha^{-1}\boldsymbol{x}) = R_{nl}(r)\, P_{lm}(\cos\vartheta) e^{im(\varphi-\alpha)} =$$

$$= \sum_{m'=-l}^{l} D_l(\alpha)_{m'm}\psi_{nlm'} = \sum_{m'=-l}^{l} D_l(\alpha)_{m'm}\, R_{nl}(r)\, P_{lm'}(\cos\vartheta)\, e^{im'\varphi}, \tag{8.4}$$

где через $D_l(\alpha)$ мы обозначили матрицу $(2l+1)$-го ранга неприводимого представления, соответствующего вращению на угол α (обозначение Вигнера от немецкого слова Darstellung — представление). Сравнивая в цепи равенств (8.4) третье звено с последним, легко видеть, что

$$D_l(\alpha)_{m'm} = \delta_{m'm} e^{-im'\alpha}, \tag{8.5}$$

где $\delta_{m'm}$ — символ Кронекера.

Мы видим, что неприводимые представления непрерывной группы вращений характеризуются числом l (размерность представления равна $2l+1$). Таким образом, s-состояние $(l=0)$, p-состояние $(l=1)$, d-состояние $(l=2)$ и т. д. являются неприводимыми представлениями электрона в сферически-симметрич-

---

[1]) Мы рассматриваем общий случай, когда $V(r)$ не кулоновская потенциальная энергия.

[2]) См., например, Блохинцев Д. И. Основы квантовой механики.— М.: Наука, 1976, § 25, 49.

ном поле. Мы видим, что как число неприводимых представлений, так и число элементов группы бесконечно.

Характер неприводимого представления $D_l(\alpha)$ равен

$$\chi_l(\alpha) = \sum_{m=-l}^{l} D_l(\alpha)_{mm} = \sum_{m=-l}^{l} e^{-im\alpha} = \\ = e^{il\alpha} + e^{i(l-1)\alpha} + \ldots + 1 + \ldots + e^{-i(l-1)\alpha} + e^{-il\alpha}.$$

Мы имеем геометрическую прогрессию из $2l+1$ членов с первым членом $e^{il\alpha}$ и знаменателем $e^{-i\alpha}$, ее сумма равна

$$\chi_l(\alpha) = \frac{e^{il\alpha}\left[\left(e^{-i\alpha}\right)^{2l+1} - 1\right]}{e^{-i\alpha} - 1} = \frac{e^{-i(l+1)\alpha} - e^{il\alpha}}{e^{-i\alpha} - 1}.$$

Умножая числитель и знаменатель на $e^{i\alpha/2}$, получим для характера неприводимого представления $D_l(\alpha)$

$$\chi_l(\alpha) = \frac{\sin(2l+1)\frac{\alpha}{2}}{\sin\frac{\alpha}{2}}. \tag{8.6}$$

2. Если поместить посторонний атом в узел кристаллической решетки (примесный атом), то сферическая симметрия поля, действующего на электрон в свободном атоме, снижается до симметрии, определяемой симметрией кристалла. Такое уменьшение симметрии поля, действующего на электрон в атоме, должно, вообще говоря, вызывать расщепление вырожденных уровней его энергии (§ 7, п. 2).

Если поле сферической симметрии заменить полем более низкой симметрии, например кубической $\boldsymbol{O}$, то базисные функции $Y_{lm}(\vartheta, \varphi)$ (8.3) и в этом случае реализуют представление группы $\boldsymbol{O}$, т. е. матрицы $D_l(R)$ (8.4), где R—элемент группы $\boldsymbol{O}$, осуществляют представление группы $\boldsymbol{O}$ размерности $(2l+1)$. Однако в общем случае это представление будет приводимым; это следует хотя бы из того, что максимальная размерность неприводимых представлений группы $\boldsymbol{O}$ равна трем, поэтому при $(2l+1) > 3$ представление $D_l(R)$ должно быть приводимым. Так как мы предполагаем, что неприводимым представлениям соответствуют разные уровни энергии атома, то, разлагая приводимое представление $D_l(R)$ на неприводимые представления группы $\boldsymbol{O}$, мы определим характер расщепления уровней энергии с разными l в поле кубической симметрии $\boldsymbol{O}$.

Воспользовавшись выражением (8.6), определим характеры для разных l, для классов группы $\boldsymbol{O}$. Например, для $l=1$

(p-состояние) имеем:

$$\chi_1(\alpha) = \frac{\sin(3\alpha/2)}{\sin(\alpha/2)},$$

$$\chi_1(E) = \chi_1(0) = \left[\frac{\sin(3\alpha/2)}{\sin(\alpha/2)}\right]_{\alpha\to 0} = 3,$$

$$\chi_1(C_3) = \chi_1(120°) = \frac{\sin 180°}{\sin 60°} = 0,$$

$$\chi_1(C_2) = \chi_1(180°) = \frac{\sin 270°}{\sin 90°} = -1,$$

$$\chi_1(C_4) = \chi_1(90°) = \frac{\sin 135°}{\sin 45°} = 1.$$

Аналогично могут быть определены характеры χ_l для других значений l. В результате мы можем составить таблицу II.9 для приводимых представлений D_l, соответствующих классам группы $\boldsymbol{O}$. Пользуясь табл. II.7 и формулой (6.34), можно определить, какие неприводимые представления кубической группы $\boldsymbol{O}$ содержатся в представлениях D_l.

$D_0 = A_1$, т. е. невырожденный уровень s-состояния не может, конечно, расщепиться и переходит в единичное представление группы $\boldsymbol{O}$.

$D_1 = F_1$, т. е. трехкратно вырожденный уровень d-состояния не расщепляется в поле кубической симметрии. (8.7)

$D_2 = E + F_2$, пятикратно вырожденное d-состояние должно расщепляться, так как наибольшая размерность неприводимых представлений группы $\boldsymbol{O}$ равна трем; d-уровень расщепляется на два уровня: двукратно вырожденный и трехкратно вырожденный.

$D_3 = A_2 + F_1 + F_2$, семикратно вырожденное f-состояние расщепляется на одно невырожденное и два трижды вырожденных состояния.

$D_4 = A_1 + E + F_1 + F_2$, девятикратно вырожденное g-состояние расщепляется на одно невырожденное, одно дважды вырожденное и два трижды вырожденных состояния.

Таблица II.9

O	E	$8C_3$	$3C_4^2$	$6C_2$	$6C_4$
D_0	1	1	1	1	1
D_1	3	0	—1	—1	1
D_2	5	—1	1	1	—1
D_3	7	1	—1	—1	—1
D_4	9	0	1	1	1

Рассмотрим теперь расщепление термов атома в поле кристалла кубической симметрии, обладающего центром инверсии J, точечной группой которого является группа $\boldsymbol{O}_h = \boldsymbol{O} \times \boldsymbol{C}_i$, где группа второго порядка $\boldsymbol{C}_i = \{E, J\}$. В этом случае при действии оператора инверсии J на волновую функцию системы последняя либо остается без изменения, либо меняет свой знак [1])

$$\hat{J}\psi(\boldsymbol{x}) = \psi(-\boldsymbol{x}) = \pm\psi(\boldsymbol{x}),$$

это непосредственно следует из того, что $J^2 = E$ (см. (7.14)) и соответственно

$$\chi(J) = \pm 1, \tag{8.8}$$

как это следует из табл. II.2.

Таким образом, характер инверсии для четных состояний системы равен $+1$, а для нечетных -1.

В первом случае волновую функцию (и соответствующее состояние) называют *четной*, а во втором — *нечетной*. Так как для рассматриваемых систем оператор $\hat{J}$ коммутирует с гамильтонианом $\hat{\mathcal{H}}$, то четность или нечетность волновой функции сохраняется во времени; это положение получило название *закона сохранения четности*. Можно показать, что для одноэлектронной волновой функции в атоме четность состояния равна $(-1)^l$, где l — азимутальное квантовое число электрона. Для многоэлектронного атома в одночастичном приближении четность равна $\prod_k (-1)^{l_k}$, где l_k — азимутальное квантовое число k-го электрона.

[1]) Ландау Л. Д., Лифшиц И. М. Квантовая механика — 3 изд. — М.: Наука, 1974, § 30.

Представим теперь, что характеры представлений $D_l(\alpha)$ определены не только для собственных вращений α, но и для несобственных $J\alpha$. Тогда табл. II.7 должна быть дополнена пятью классами: $JE=J$, $8JC_3$, $3JC_4^2$ и т. д. Характеры представлений этих дополнительных классов либо совпадают с характерами табл. II.7 (для четных состояний), либо имеют противоположные знаки (для нечетных состояний).

Так как в табл. II.3 представление Γ_1 соответствует четным, а Γ_2—нечетным состояниям, то таблица характеров неприводимых представлений группы (включающей инверсию J) имеет вид, изображенный на табл. II.6, где χ совпадает с табл. II.7.

Рассмотрим теперь, что вносят отдельные квадранты табл. II.6 в сумму при вычислении коэффициентов a_j по формуле (6.34). Квадранты 3 и 4 в сумме сокращаются, а 1 и 2 дают одинаковые вклады, которые затем делятся на вдвое большее число h. В результате для a_j получается то же значение, что и для одного квадранта 1, т. е. как и для группы $\boldsymbol{O}$.

3. Покажем теперь, как применять теорию групп к классификации нормальных колебаний многоатомных молекул. Мы рассмотрим задачу о малых колебаниях атомных ядер молекулы только в рамках классической механики.

Как известно из классической механики, система из N частиц, обладающая $3N$ степенями свободы и совершающая малые колебания около положения равновесия частиц $\boldsymbol{r}_i$ $(i=1, 2, \ldots, N)$ обладает энергией

$$\mathscr{E}=\frac{1}{2}\sum_{i,\,\alpha}\frac{\dot{u}_{i\alpha}^2}{m_i}+\frac{1}{2}\sum_{i,\,k,\,\alpha,\,\beta}\varkappa_{i\alpha,\,k\beta}u_{i\alpha}u_{k\beta}. \tag{8.9}$$

Здесь $u_{i\alpha}$ $(i=1, 2, \ldots, N;\ \alpha=x, y, z)$—$\alpha$-я прямоугольная (декартова) составляющая отклонения i-го атомного ядра молекулы от положения равновесия $\boldsymbol{r}_i$, $\varkappa_{i\alpha,\,k\beta}$—коэффициент квазиупругой силы для отклонений $u_{i\alpha}$ и $u_{k\beta}$. Из общего числа $3N$ степеней свободы молекулы три соответствуют ее поступательному смещению как целому, три другие—вращательному движению молекулы как целому. Таким образом, собственно колебательных степеней свободы в молекуле $3N-6$. Соответствующим линейным преобразованием можно из (8.9) исключить эти шесть степеней свободы. Затем посредством соответствующего линейного преобразования можно ввести *нормальные координаты* $Q_{i\alpha}$ $(i,\ \alpha=1, 2, 3, \ldots, (3N-6)$ [1]); тогда энергии $\mathscr{E}$ приобретают вид

$$\mathscr{E}=\frac{1}{2}\sum_{i,\,\alpha}\dot{Q}_{i\alpha}^2+\frac{1}{2}\sum_i\omega_i^2\sum_{\alpha=1}^{f_i}Q_{i\alpha}^2, \tag{8.10}$$

[1]) Ландау Л. Д., Лифшиц Е. М. / Теоретическая физика, т. 1.—Механика.—4 изд.—М.: Наука, 1973, гл. V.

где ω_i — частота нормального колебания. Мы ввели у нормальных координат $Q_{i\alpha}$ двойной индекс $i\alpha$, где i пробегает значения, соответствующие различным частотам ω_i, а α пробегает f_i значений соответствующих f_i линейно независимым нормальным координатам колеблющимся с данной частотой ω_i; величина f_i называется кратностью вырождения частоты ω_i; очевидно, что $\sum_i f_i = 3N-6$.

Обычно структура молекулы удовлетворяет некоторой группе точечной симметрии. Например, молекула аммиака NH_3 имеет форму правильной треугольной пирамиды, в вершине которой расположен атом азота N, а в вершинах равностороннего треугольника основания три атома водорода H. Молекула NH_3 обладает вертикальной осью симметрии C_3, проходящей через атом N и тремя плоскостями σ_v, каждая из которых проходит через ось C_3 и один атом H. Группа симметрии молекулы аммиака $\boldsymbol{C_{3v}}$ состоит из шести элементов $E, 2C_3, 3\sigma_v$. При преобразованиях группы симметрии молекулы ее динамические коэффициенты $\varkappa_{i\alpha,\ k\beta}$ и m_i не меняются, поэтому не меняются нормальные частоты колебаний ω_i.

Подействуем оператором $\hat{P}_R$ (6.1), где R — элемент симметрии молекулы, на потенциальную энергию в (8.10):

$$\hat{P}_R\left\{\frac{1}{2}\sum_i \omega_i^2 \sum_{\alpha=1}^{fi} Q_{i\alpha}^2\right\} = \frac{1}{2}\sum_i \omega_i^2 \sum_{\alpha=1}^{fi} \hat{P}_R Q_{i\alpha}^2 =$$

$$= \frac{1}{2}\sum_i \omega_i^2 \sum_{\alpha=1}^{fi} (\hat{P}_R Q_{i\alpha})(\hat{P}_R Q_{i\alpha}). \quad (8.11)$$

Мы видим, что для того, чтобы преобразованные нормальные координаты $\hat{P}_R Q_{i\alpha}$ сохранили смысл нормальных координат при тех же частотах ω_i, необходимо, чтобы

$$\hat{P}_R Q_{i\alpha} = \sum_\beta \Gamma(R)^{(i)}_{\beta\alpha} Q_{R^{-1}(i)\,\beta}, \quad (8.11a)$$

где $R^{-1}(i) = k$ — положение идентичного атомного ядра молекулы, а $\Gamma(R)^{(i)}_{\beta\alpha}$ — матрица линейного ортогонального преобразования. Совершенно так же, как это было сделано в § 6, можно показать, что матрицы $\Gamma(R)^{(i)}$ образуют представление группы симметрии молекулы с базисом $Q_{i\alpha}$ ($\alpha = 1, 2, \ldots, f_i$). Если исключить случайное вырождение частот ω_i — это представление будет неприводимым. Размерность этого неприводимого представления равна кратности вырождения частоты колебания f_i.

Преобразование (8.11а) аналогично преобразованию (7.15) в квантовой механике. Роль вырожденного уровня энергии $\mathscr{E}$ играет частота нормальных колебаний ω_i, а роль волновых

функций $\Psi_i(\boldsymbol{x})$ $(i=1, 2, \ldots, l)$ — нормальные колебания $Q_{i\alpha}$ $(\alpha=1, 2, \ldots, f_i)$.

Задача определения собственных частот и нормальных координат для многоатомной молекулы — сложная динамическая задача. Теория групп позволяет, не решая этой задачи, определить число собственных частот ω_i, кратность их вырождения и даже форму соответствующих им колебаниям.

Для выполнения этой программы мы определим общее приводимое представление, осуществляемое сразу всеми колебательными координатами, а затем разложим его по неприводимым представлениям группы симметрии молекулы. Для нахождения полного представления используем то, что характеры представлений инвариантны относительно преобразования подобия. Поэтому для их вычисления можно воспользоваться не нормальными координатами, а прямоугольными смещениями ядер $u_{i\alpha}$.

Определим действие оператора $\hat{P}_R$ на поле смещений $u_{i\alpha}$, когда R — элемент точечной группы симметрии молекулы, перемещающий ядро i в тождественное ядро $k=R^{-1}(i)$.

На рис. II.17 представлена плоская молекула из четырех одинаковых атомов *1*, *2*, *3*, *4*, расположенных на концах равностороннего прямоугольного креста. Пусть $\hat{R}^{-1}$ вращение молекулы вокруг ее центра O по часовой стрелке на угол $\pi/2$. Атом *1* совмещается при этом с атомом *2*, атом *2* — с атомом *3* и т. д. Если бы мы переместили атомы *1*, *2* и т. д. вместе с их смещениями $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ и т. д., то никакой новой информации мы не получили бы, так как это эквивалентно повороту молекулы как целой в пространстве. Мы можем извлечь что-либо из симметрии молекулы, если перенесем смещение $\boldsymbol{u}_1$ на атом *2*: смещение $\boldsymbol{u}_2$ на атом *3* и т. д. Такое преобразование поля смещений $\boldsymbol{u}_i$ будет ввиду одинаковости атомов соответствовать тем же частотам ω_i колебаний молекулы.

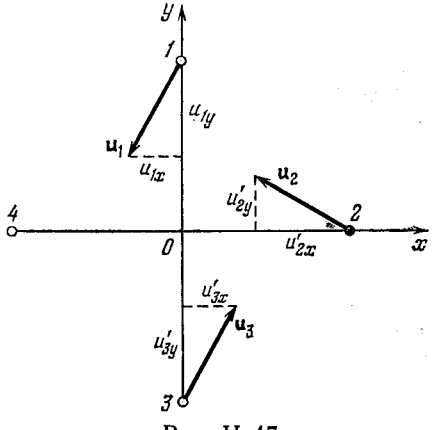


Рис. II.17.

Как преобразуются при повороте C_4 смещения u_{1x} и u_{1y}? Из рис. II.17 видно, что

$$\hat{P}_{C_4^{-1}}u_{1x}=u'_{2y}=-u_{1x}, \quad \hat{P}_{C_4^{-1}}u_{1y}=u'_{2x}=u_{1y}. \tag{8.12}$$

Если R^{-1} — вращение на угол π, так что атом *1* переходит

в атом 3, то

$$P_{C_2^{-1}} u_{1x} = u'_{3x} = -u_{1x}, \quad P_{C_2^{-1}} u_{1y} = u'_{3y} = -u_{1y}. \tag{8.12a}$$

Обобщая (8.12) и (8.12a) на более сложную конфигурацию молекулы, выразим $\hat{P}_R u_{i\alpha}$ в виде линейной комбинации величин $u_{k\beta}$, т. е.

$$\hat{P}_R u_{i\alpha} = \sum_{k\beta}^{1,\,3N} D^u(R)_{k\beta,\,i\alpha} u_{k\beta} = \sum_{k\beta} A(R)_{\beta\alpha} \delta_{k,R^{-1}(i)} u_{k\beta}, \tag{8.13}$$

где $R^{-1}(i) = k$, т. е. ядро i переходит, при преобразовании R^{-1}, в идентичное ядро k. Матрица D^u $3N$-го ранга дает приводимое представление, осуществляемое всеми колебаниями молекулы.

Так как операции R^{-1} смещения $u_{i\alpha}$ преобразуются как компоненты полярного вектора (8.12), то $D^u(R)_{k\beta,\,i\alpha} = A(R)_{\beta\alpha}\delta_{k,R^{-1}(i)}$, где $A(R)$ ортогональная матрица третьего ранга.

Нас интересует характер (шпур) матрицы $D^u(R)$, поэтому, если $k \neq i$, т. е. преобразование R^{-1} перемещает ядро i в д р у г о е идентичное ядро k, то соответствующий элемент матрицы $D^u(R)$ не лежит на главной диагонали и вклада в характер не дает. Поэтому нас интересуют только те ядра, которые при преобразовании R^{-1} остаются на месте, но для них матрица D^u приобретает квазидиагональный вид, все блоки которой (число блоков равно числу атомов, остающихся на месте при преобразовании R^{-1}) равны $A(R)$.

Если преобразование R^{-1} есть вращение смещения $\boldsymbol{u}$ (мы опускаем индекс i) на угол φ вокруг оси z, то

$$\hat{P}_{C(\varphi)} u_\alpha = \sum_{\beta=1}^{3} A(C)_{\beta\alpha} u_\beta = \begin{cases} u_x \cos\varphi + u_y \sin\varphi, \\ -u_x \sin\varphi + u_y \cos\varphi, \\ u_z. \end{cases} \tag{8.14}$$

Шпур матрицы $A(C)$ равен $1 + 2\cos\varphi$. Если на оси вращения расположено N_C ядер, то шпур матрицы D^u в (8.13) равен

$$N_C(1 + 2\cos\varphi). \tag{8.15}$$

Однако этот шпур связан со всеми $3N$ степенями свободы молекулы, в том числе и с поступательными и вращательными перемещениями молекулы как целого. Поступательное перемещение молекулы определяется полярным вектором смещения центра тяжести $\boldsymbol{U}$, для которого характер (шпур) преобразования, так же как и для колебаний $\boldsymbol{u}$, равен $(1 + 2\cos\varphi)$. Поворот молекулы как целого на малый угол $\delta\Omega$ характеризуется аксиальным вектором $\boldsymbol{\delta\Omega}$, направленным по оси вращения (по правилу буравчика). При собственных вращениях координатной системы аксиальный вектор $\boldsymbol{\delta\Omega}$ ведет себя как полярный, которому соответствует шпур, так же равный $(1 + 2\cos\varphi)$. Таким образом, для определения

характера полного представления, соответствующего колебательным степеням свободы, надо из шпура (8.15) вычесть $2(1+2\cos\varphi)$, так что

$$\chi^u(C)=(N_C-2)(1+2\cos\varphi). \tag{8.16}$$

Характер единичного элемента ($N_C=N$, $\varphi=0$) равен $\chi^u(E)=(N-2)\cdot 3=3N-6$. Аналогично вычисляется характер полного колебательного представления для зеркально-поворотного преобразования $S(\varphi)$, т. е. для вращения на угол φ вокруг оси z и отражения в плоскости xy. Если ядро атома не смещается при этом преобразовании, то смещение преобразуется, аналогично (8.14), по формулам

$$\hat{P}_{S(\varphi)}u_\alpha=\sum_{\beta=1}^{3}A(S)_{\beta\alpha}u_\beta=\begin{cases}u_x\cos\varphi+u_y\sin\varphi,\\ -u_x\sin\varphi+u_y\cos\varphi,\\ -u_z,\end{cases} \tag{8.17}$$

которым соответствует шпур $(-1+2\cos\varphi)$. Характер представления, осуществляемого всеми $3N$ степенями свободы, равен

$$N_S(-1+2\cos\varphi), \tag{8.18}$$

где N_S — число ядер, неподвижных при операции $S(\varphi)$. Вектору смещения центра тяжести $\boldsymbol{U}$ соответствует при операции $S(\varphi)$ шпур $(-1+2\cos\varphi)$, а аксиальному вектору $\delta\boldsymbol{\Omega}$ при той же операции — шпур $(1-2\cos\varphi)$; последнее следует из того, что при отражении в плоскости xy правая координатная система переходит в левую, а при этом аксиальный вектор $\delta\boldsymbol{\Omega}$ меняет направление на противоположное; т. е. в правой части преобразования (8.17) надо заменить u_α на $-u_\alpha$. Мы видим, что суммарный шпур от $\boldsymbol{U}$ и $\delta\boldsymbol{\Omega}$ при операции $S(\varphi)$ равен $-1+2\cos\varphi+1-2\cos\varphi=0$, поэтому (8.18) представляет собой характер полного представления всех колебательных степеней свободы для операции $S(\varphi)$:

$$\chi^u(S)=N_S(-1+2\cos\varphi). \tag{8.19}$$

Если $S(\varphi)$ сводится к одному отражению, т. е. $S=\sigma$ $(\varphi=0)$, то $\chi^u(\sigma)=N_S$.

Для классификации нормальных колебаний молекулы достаточно теперь полное представление (8.16), (8.19) разложить по неприводимым представлениям точечной группы симметрии молекулы.

Рассмотрим молекулу аммиака NH_3, о которой упоминалось выше. Как уже отмечалось, точечная группа ее симметрии $\boldsymbol{C_{3v}}$ состоит из шести элементов, распределенных по трем классам: E, $2C_3$, $3\sigma_v$ (см. § 3, группа IV C_{nv}).

Определим, пользуясь формулами (8.16) и (8.19), характеры полного представления D^a для молекулы NH_3:

$$\chi^a(E) = (N-2)(1+2\cos 0^\circ) = (N-2)\cdot 3 = (4-2)\cdot 3 = 6,$$
$$\chi^a(C_3) = (N_C-2)(1+2\cos 120^\circ) = -(1+2\cos 120^\circ) = 0, \quad (8.20)$$
$$\chi^a(\sigma) = N_6 = 2.$$

Пользуясь формулами (6.34), (8.20) и табл. II.3 характеров группы $\boldsymbol{D}_3$, изоморфной с группой $\boldsymbol{C}_{3v}$, легко показать, что приводимое представление D^a распадается на следующие неприводимые представления группы $\boldsymbol{C}_{3v}$:

$$D^a = 2\Gamma_1 + 2\Gamma_3 = 2A_1 + 2E, \quad (8.21)$$

где мы, как это принято в молекулярной физике, обозначили единичное представление Γ_1 через A_1 и дважды вырожденное представление Γ_3 через E. Таким образом, молекула аммиака NH_3, имеющая шесть колебательных степеней свободы, обладает двумя невырожденными нормальными колебаниями типа A_1 и двумя дважды вырожденными колебаниями типа E. При колебаниях типа A_1, соответствующих единичному представлению, полная симметрия молекулы (правильная треугольная пирамида) сохраняется. Мы не будем заниматься рассмотрением характера смещений атомов для колебаний типа E.

§ 9. Применение теории групп к трансляционной симметрии кристалла

1. Важнейшим видом симметрии твердого тела является его трансляционная симметрия, согласно которой кристаллическая решетка совпадает сама с собой при смещении кристалла как целого на вектор решетки

$$\boldsymbol{a}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3, \quad (9.1)$$

где $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$ — основные векторы, а n_i — целые числа (гл. I, § 3, п. 1). Элементы трансляционной группы $\mathscr{T}$, являющейся инвариантной подгруппой пространственной группы кристалла, могут быть записаны в виде (5.4)

$$t_n = \{E \mid \boldsymbol{a}_n\} = \{E \mid n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3\} =$$
$$= \{E \mid n_1\boldsymbol{a}_1\}\{E \mid n_2\boldsymbol{a}_2\}\{E \mid n_3\boldsymbol{a}_3\} = t_{n_1}t_{n_2}t_{n_3}, \quad (9.2)$$

где $R = E$ — единичный элемент точечной группы кристалла.

Из (9.2) видно, что группа $\mathscr{T}$ может быть представлена как прямое произведение трех одномерных трансляционных групп с элементами $t_{n_i} = \{E \mid n_i\boldsymbol{a}_i\}$ $(i = 1, 2, 3)$.

Группа трансляций — абелева (и циклическая), поэтому каждый элемент ее является классом, и ее неприводимые представ-

ления имеют размерность единица, т. е. являются числами (вообще говоря, комплексными). В самом деле, для абелевых групп число классов s, равное числу неприводимых представлений, равно порядку группы h; тогда из (6.22) следует, что размерности всех неприводимых представлений $l_i = 1$.

Для того чтобы избежать трудностей, связанных для конечного кристалла с заданием граничных условий, разобьем бесконечный кристалл на одинаковые параллелепипеды с ребрами $G\boldsymbol{a}_1$, $G\boldsymbol{a}_2$, $G\boldsymbol{a}_3$, где G — большое число (удобно считать G нечетным числом). Мы заменим граничные условия условиями *цикличности Борна—Кармана*, согласно которым все физические свойства и функции (волновая функция, волна колебаний атомов кристалла и т. д.) имеют одинаковое значение в точках $\boldsymbol{r}$ и $\boldsymbol{r}+G\boldsymbol{a}_i$ $(i=1, 2, 3)$, т. е. периодически повторяются во всех параллелепипедах, на которые мы разбили бесконечный кристалл.

Таким образом, применение условий цикличности позволяет рассматривать все явления и свойства кристалла в пределах одного выделенного в нем параллелепипеда (*основной области*) объема $V = G^3\Omega_0$, где $\Omega_0 = (\boldsymbol{a}_1[\boldsymbol{a}_2, \boldsymbol{a}_3])$ — объем элементарной ячейки кристалла.

Условиям цикличности Борна—Кармана в случае трехмерного кристалла нельзя придать столь наглядный смысл, как в случае одномерной цепочки равноотстоящих атомов (см. гл. III, § 2, п. 3). Однако можно показать, что в математическом отношении условия цикличности для трехмерного кристалла эквивалентны любым граничным условиям на поверхности основной области, не влияющим на объемные свойства кристалла [1]).

Условия цикличности делают циклическую трансляционную группу конечной, так как смещению $G\boldsymbol{a}_i$ соответствует единичный элемент смещения $\{E\,|\,G\boldsymbol{a}_i\} = \{E\,|\,0\}$; таким образом, трансляционная группа $\mathscr{T}$ имеет порядок G^3.

Рассмотрим неприводимые представления группы трансляций $\mathscr{T}$. Определим вначале неприводимые представления группы одномерной трансляции вдоль вектора $\boldsymbol{a}_1$, состоящей из элементов $t_{n_1} = \{E\,|\,n_1\boldsymbol{a}_1\}$.

Пусть элементу $\{E\,|\,\boldsymbol{a}_1\}$ соответствует неприводимое представление (число) s, что мы обозначим следующим образом:

$$\{E\,|\,\boldsymbol{a}_1\} \longrightarrow s,$$

тогда получим для элементов

$$\begin{aligned} \{E\,|\,2\boldsymbol{a}_1\} &= \{E\,|\,\boldsymbol{a}_1\}\{E\,|\,\boldsymbol{a}_1\} \longrightarrow s\cdot s = s^2, \\ \{E\,|\,n_1\boldsymbol{a}_1\} &= \{E\,|\,\boldsymbol{a}_1\}\ldots\{E\,|\,\boldsymbol{a}_1\} \longrightarrow s\cdot s\cdot\ldots\cdot s = s^{n_1}, \\ \{E\,|\,G\boldsymbol{a}_1\} &= \{E\,|\,0\} \longrightarrow s^G = 1. \end{aligned} \qquad (9.3)$$

[1]) Борн М., Кунь Х. Динамическая теория кристаллических решеток.—М.: ИЛ, 1958, с. 453.

Последнее следует из того, что единичному элементу смещения $\{E\,|\,0\}$ соответствует неприводимое представление (число) — единица. Из последнего равенства следует [1])

$$s=\sqrt[G]{1}=e^{\frac{2\pi i}{G}g_1}, \tag{9.4}$$

где

$$g_1=0,\ 1,\ 2,\ \ldots,\ G-1.$$

Из (9.3) и (9.4) следует, что элементу $t_{n_1}=\{E\,|\,n_1\boldsymbol{a}_1\}$ соответствуют G неприводимых представлений

$$\Gamma_{g_1}(t_{n_1})=s^{n_1}=\exp\frac{2\pi i}{G}g_1n_1,\quad g_1=0,\ 1,\ 2,\ \ldots,\ G-1. \tag{9.5}$$

Таким образом, в согласии с общей теорией, число неприводимых представлений равно числу классов (элементов) (см. (6.30)). Аналогично для двух других групп одномерной трансляции с элементами $t_{n_2}=\{E\,|\,n_2\boldsymbol{a}_2\}$ и $t_{n_3}=\{E\,|\,n_3\boldsymbol{a}_3\}$, неприводимые представления равны

$$\begin{aligned}\Gamma_{g_2}(t_{n_2})&=\exp\frac{2\pi i}{G}g_2n_2,\quad g_2=0,\ 1,\ 2,\ \ldots,\ G-1,\\ \Gamma_{g_3}(t_{n_3})&=\exp\frac{2\pi i}{G}g_3n_3,\quad g_3=0,\ 1,\ 2,\ \ldots,\ G-1.\end{aligned} \tag{9.6}$$

В § 6, п. 5 было показано, что неприводимое представление прямого произведения групп равно прямому произведению их неприводимых представлений. Как уже отмечалось выше, трехмерная группа трансляций $\mathscr{T}$ равна прямому произведению трех одномерных групп трансляции с элементами t_{n_1}, t_{n_2} и t_{n_3}, поэтому неприводимое представление группы $\mathscr{T}$ равно

$$\Gamma_{g_1g_2g_3}(t_n)=\Gamma_{g_1}(t_{n_1})\,\Gamma_{g_2}(t_{n_2})\,\Gamma_{g_3}(t_{n_3})=$$
$$=\exp\frac{2\pi i}{G}(g_1n_1+g_2n_2+g_3n_3), \tag{9.7}$$

как это следует из (9.5) и (9.6).

Введем *волновой вектор*

$$\boldsymbol{k}=\frac{g_1}{G}\boldsymbol{b}_1+\frac{g_2}{G}\boldsymbol{b}_2+\frac{g_3}{G}\boldsymbol{b}_3, \tag{9.8}$$

где $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ и $\boldsymbol{b}_3$ — основные векторы обратной решетки (I.3.8). Очевидно, что $\boldsymbol{k}$ имеет размерность векторов $\boldsymbol{b}_i$, т. е. размерность, обратную длине ($см^{-1}$). Тройка чисел g_1, g_2, g_3 определяет вектор $\boldsymbol{k}$. Из (I.3.9) следует, что показатель экспоненты (9.7) может быть представлен в виде $i\boldsymbol{k}\boldsymbol{a}_n$, поэтому (9.7) может быть записано в виде

$$\Gamma_{\boldsymbol{k}}(t_n)=e^{i\boldsymbol{k}\boldsymbol{a}_n}, \tag{9.9}$$

---

[1]) Смирнов В. И. Курс высшей математики.—23 изд.—М.: Наука, 1974, т. 1, § 175.

где волновой вектор $\boldsymbol{k}$ нумерует неприводимое представление. Если заменить вектор $\boldsymbol{k}$ на $\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{b}_m$, где $\boldsymbol{b}_m = m_1\boldsymbol{b}_1 + m_2\boldsymbol{b}_2 + m_3\boldsymbol{b}_3$ — произвольный вектор обратной решетки, то

$$\exp(i\boldsymbol{k}'\boldsymbol{a}_n) = \exp(i\boldsymbol{k}\boldsymbol{a}_n)\exp(i\boldsymbol{b}_m\boldsymbol{a}_n) = \\ = \exp(i\boldsymbol{k}\boldsymbol{a}_n)\exp(2\pi i \times \text{целое число}) = \exp(i\boldsymbol{k}\boldsymbol{a}_n),$$

где мы воспользовались формулой (I.3.9).

Таким образом, $\boldsymbol{k}$ и $\boldsymbol{k}'$ соответствуют одному и тому же неприводимому представлению, т. е. они эквивалентны. В силу этого можно ограничиться рассмотрением таких значений $\boldsymbol{k}$, для которых $\boldsymbol{k}\boldsymbol{a}_i$ $(i = 1, 2, 3)$ лежат внутри интервала 2π (так как, если $\boldsymbol{k}\boldsymbol{a}_i$ лежит вне этого интервала, то оно может быть приведено к нему исключением из $\boldsymbol{k}$ соответствующего вектора обратной решетки); мы можем, например, положить

$$-\pi < \boldsymbol{k}\boldsymbol{a}_i < \pi \quad (i = 1, 2, 3). \tag{9.10}$$

Подставляя сюда вместо $\boldsymbol{k}$ его значение (9.8), получим

$$-G/2 < g_i < G/2, \tag{9.10a}$$

т. е. мы приходим к тем же условиям (9.5), (9.6), согласно которым g_i пробегает G значений (из (9.10а) видно, почему удобно выбирать G нечетным числом).

В качестве области неэквивалентных значений вектора (9.10) может быть выбрана элементарная ячейка обратной решетки «объема» $(2\pi)^3/\Omega_0$. Такой выбор области (9.10) не всегда удобен, так как параллелепипед, построенный на векторах $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$ не обладает, вообще говоря, симметрией самой решетки (прямой или обратной). Аналогично тому, как была построена симметричная ячейка Вигнера—Зейтца (гл. I, § 1, п. 4), можно и в случае обратной решетки всегда выделить область (9.10), обладающую полной симметрией обратной (и прямой) решетки. Для этого проведем из некоторого узла O обратной решетки векторы $\boldsymbol{b}_g$ ко всем другим ее узлам. Проведем плоскости через середины векторов $\boldsymbol{b}_g$, перпендикулярные к ним. Уравнения этих плоскостей имеют вид

$$^1/_2\boldsymbol{b}_g^2 - (\boldsymbol{b}_g\boldsymbol{k}) = 0, \tag{9.11}$$

где $\boldsymbol{k}$ — радиус-вектор, проведенный из начала координат O к некоторой точке, построенной плоскости. Заметим, что уравнение (9.11) совпадает с условиями дифракции рентгеновских лучей в кристалле (I.4.5) (разные знаки перед скалярным произведением в (I.4.5) и (9.11) связаны с тем, что $\boldsymbol{k}$ в уравнении (I.4.5) направлен от плоскости к началу координат O, см.

рис. I.16). Эти пересекающиеся плоскости определяют некоторые многогранники с центром в точке O. Наименьший многогранник с центром в точке O называется *первой* или *приведенной зоной Бриллюэна* (иногда просто *бриллюэновской зоной*); его «объем» тоже равен $(2\pi)^3/\Omega_0$ (см. ниже). «Объем», заключенный между поверхностью первой бриллюэновской зоны и поверхностью следующего многогранника, называется *второй бриллюэновской зоной* и т. д. Можно показать, что «объемы» всех бриллюэновских зон одинаковы и равны $(2\pi)^3/\Omega_0$.

Как уже отмечалось в (гл. I, § 3), обратная решетка для простой кубической решетки тоже простая кубическая; отсюда сразу следует, что первая бриллюэновская зона в этом случае имеет форму куба с ребром $2\pi/a$ (a — ребро куба прямой решетки); «объем» зоны Бриллюэна равен $(2\pi/a)^3 = (2\pi)^3/a^3 = (2\pi)^3/\Omega_0$.

Из самого построения первой зоны Бриллюэна следует, что волновые векторы $\boldsymbol{k}$, концы которых лежат внутри нее, отличаются друг от друга меньше чем на вектор обратной решетки. Если же конец вектора $\boldsymbol{k}$ лежит на границе зоны Бриллюэна, то всегда существует по крайней мере один эквивалентный ему вектор $\boldsymbol{k}' = \boldsymbol{k} \pm \boldsymbol{b}_i$, конец которого также лежит на границе бриллюэновской зоны.

Первая зона Бриллюэна может быть определена и как множество точек, расстояние которых до данного узла обратной решетки меньше, чем расстояние до всех других узлов (только точки, лежащие на границе зоны, отстоят на одинаковых расстояниях от двух узлов обратной решетки). Таким образом, все пространство обратной решетки можно разбить на приведенные бриллюэновские зоны, построенные около каждого из узлов.

Рассмотрим обратную решетку, соответствующую основной области кристалла, содержащей G^3 элементарных ячеек. Полный объем такой обратной решетки равен объему элементарной ячейки обратной решетки $(2\pi)^3/\Omega_0$, помноженной на G^3. С другой стороны, этот объем равен объему бриллюэновской зоны, помноженной на G^3; отсюда следует, что объем бриллюэновской зоны равен объему элементарной ячейки обратной решетки, т. е. равен $(2\pi)^3/\Omega_0$, как и утверждалось выше.

Более подробное исследование бриллюэновских зон будет дано в гл. IV при изучении движения электрона в идеальном кристалле.

Мы ввели волновой вектор $\boldsymbol{k}$, исходя только из свойств трансляционной симметрии. И хотя, например, в следующей главе волновой вектор появится при решении механической задачи о колебании атомов кристалла, мы должны отдавать себе отчет в том, что более глубокая причина его появления — трансляционная симметрия кристалла.

2. Рассмотрим действие оператора $\hat{P}_R$ (6.1), где R — трансляция на вектор $-\boldsymbol{a}_n$ [1]) на волновую функцию электрона $\psi_{\boldsymbol{k}}(\boldsymbol{r})$, движущегося в периодическом поле кристалла. Индекс $\boldsymbol{k}$ у волновой функции означает, что мы рассматриваем базисную функцию $\boldsymbol{k}$-го неприводимого представления.

Так как неприводимые представления группы трансляции имеют размерность единица и равны (9.9), то из (6.1) следует

$$\hat{P}_R\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \psi_{\boldsymbol{k}}(R^{-1}\boldsymbol{r}) = \psi_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{a}_n) = \Gamma_{\boldsymbol{k}}\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\boldsymbol{a}_n}\psi_{\boldsymbol{k}}(\boldsymbol{r}),$$

т. е.

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{a}_n) = e^{i\boldsymbol{k}\boldsymbol{a}_n}\psi_{\boldsymbol{k}}(\boldsymbol{r}). \tag{9.12}$$

Легко показать, что для того, чтобы волновая функция электрона $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ удовлетворяла трансформационному соотношению (9.12), необходимо, чтобы

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = u_{\boldsymbol{k}}(\boldsymbol{r})\, e^{i\boldsymbol{k}\boldsymbol{r}}, \tag{9.13}$$

где функция $u_{\boldsymbol{k}}(\boldsymbol{r})$ обладает периодичностью решетки, т. е.

$$u_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{a}_n) = u_{\boldsymbol{k}}(\boldsymbol{r}). \tag{9.13а}$$

В самом деле, в этом случае из (9.13) следует

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{a}_n) = u_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{a}_n)\, e^{i\boldsymbol{k}(\boldsymbol{r}+\boldsymbol{a}_n)} = u_{\boldsymbol{k}}(\boldsymbol{r})\, e^{i\boldsymbol{k}\boldsymbol{r}} e^{i\boldsymbol{k}\boldsymbol{a}_n} = e^{i\boldsymbol{k}\boldsymbol{a}_n}\psi_{\boldsymbol{k}}(\boldsymbol{r}),$$

что совпадает с (9.12).

Волновая функция электрона в кристалле (9.13) называется *блоховской волновой функцией* (Ф. Блох, 1928 г.); ее вид обусловлен только трансляционной симметрией кристалла. Конкретный вид функции $u_{\boldsymbol{k}}(\boldsymbol{r})$ зависит от вида периодического потенциала, действующего на электрон.

3. Стационарные состояния электрона в периодическом поле кристалла описываются уравнением Шредингера

$$\hat{\mathcal{H}}(\boldsymbol{r})\,\psi_{n\boldsymbol{k}j}(\boldsymbol{r}) \equiv \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})\right]\psi_{n\boldsymbol{k}j}(\boldsymbol{r}) = \mathcal{E}_n(\boldsymbol{k})\,\psi_{n\boldsymbol{k}j}(\boldsymbol{r}), \tag{9.14}$$

где гамильтониан $\hat{\mathcal{H}}(\boldsymbol{r})$ инвариантен при всех преобразованиях симметрии кристалла.

Блоховская волновая функция электрона (9.13)

$$\psi_{n\boldsymbol{k}j}(\boldsymbol{r}) = u_{n\boldsymbol{k}j}(\boldsymbol{r})\, e^{i\boldsymbol{k}\boldsymbol{r}} \tag{9.15}$$

зависит не только от волнового вектора $\boldsymbol{k}$, но и от номера полосы (зоны) энергии n; в общем случае она может зависеть еще от дополнительных квантовых чисел j, характеризующих разные вырожденные состояния электрона при заданных n и $\boldsymbol{k}$.

[1]) Можно было бы, конечно, определить операцию R как трансляцию на вектор $+\boldsymbol{a}_n$, но это привело бы к менее привычным соотношениям.

В (9.14) $\mathscr{E}_n(\boldsymbol{k})$ — собственные значения энергии электрона. Пусть группой симметрии кристалла является пространственная группа с элементами (§ 5 п. 3)

$$g=\{R\,|\,\boldsymbol{\alpha}(R)+\boldsymbol{a}_n\} \tag{9.16}$$

и с обратными элементами (5.6)

$$g^{-1}=\{R^{-1}\,|\,-R^{-1}\boldsymbol{\alpha}(R)-R^{-1}\boldsymbol{a}_n\}. \tag{9.17}$$

Подействуем оператором $\hat{P}_g$ (6.1) на обе части уравнения (9.14); так как периодическое поле кристалла $V(\boldsymbol{r})$ и оператор кинетической энергии $(-\hbar^2/2m)\nabla^2$ не меняются при всех преобразованиях пространственной группы (9.16), то

$$\hat{\mathscr{H}}(\boldsymbol{r})\,\hat{P}_g\psi_{nkj}(\boldsymbol{r})=\mathscr{E}_n(\boldsymbol{k})\,\hat{P}_g\psi_{nkj}(\boldsymbol{r}), \tag{9.18}$$

т. е. функция $\hat{P}_g\psi_{nkj}(\boldsymbol{r})$ тоже должна быть блоховской функцией, соответствующей той же энергии $\mathscr{E}_n(\boldsymbol{k})$.

Из (6.1), (9.17) и (9.15) следует

$$\hat{P}_g\psi_{nkj}(\boldsymbol{r})=\hat{P}_g[u_{nkj}(\boldsymbol{r})\,e^{i\boldsymbol{k}\boldsymbol{r}}]=u_{nkj}(g^{-1}\boldsymbol{r})\,e^{ikg^{-1}r}=u_{nkj}(R^{-1}\boldsymbol{r}-$$
$$-R^{-1}\boldsymbol{\alpha}(R)-R^{-1}\boldsymbol{a}_n)\exp[-i\boldsymbol{k}(R^{-1}\boldsymbol{\alpha}+R^{-1}\boldsymbol{a}_n)]\,e^{ikR^{-1}r}. \tag{9.19}$$

Вектор $R^{-1}\boldsymbol{a}_n$ равен вектору решетки (это справедливо и для несимморфных групп, когда R^{-1} может и не быть элементом симметрии сложной решетки). Если учесть еще, что скалярное произведение векторов не изменится, если применить к каждому из них преобразование R, т. е.

$$\boldsymbol{k}\cdot R^{-1}\boldsymbol{r}=R\boldsymbol{k}\cdot RR^{-1}\boldsymbol{r}=R\boldsymbol{k}\cdot E\boldsymbol{r}=R\boldsymbol{k}\boldsymbol{r}, \tag{9.20}$$

то (9.19) может быть записано в виде

$$\hat{P}_g\psi_{nkj}(\boldsymbol{r})=\tilde{u}_{nRkj}(\boldsymbol{r})\,e^{iR\boldsymbol{k}\boldsymbol{r}}, \tag{9.21}$$

где

$$\tilde{u}_{nRkj}(\boldsymbol{r})=u_{nkj}(R^{-1}\boldsymbol{r}-R^{-1}\boldsymbol{\alpha}(R))\exp[-iR\boldsymbol{k}(\boldsymbol{\alpha}(R)+\boldsymbol{a}_n)] \tag{9.21a}$$

— периодическая функция с периодами решетки (это следует из того, что $R^{-1}\boldsymbol{a}_m$ равен вектору решетки). Блоховская функция (9.21) с волновыми векторами $R\boldsymbol{k}$ удовлетворяет тому же уравнению Шредингера (9.14) с собственным значением энергии $\mathscr{E}_n(\boldsymbol{k})$; другими словами,

$$\mathscr{E}_n(\boldsymbol{k})=\mathscr{E}_n(R\boldsymbol{k}). \tag{9.22}$$

Таким образом, энергетические поверхности $\mathscr{E}_n(\boldsymbol{k})=\text{const}$ в бриллюэновской зоне для всех полос (зон) энергии обладают симметрией точечной группы кристалла.

Совокупность векторов $R\boldsymbol{k}$, где R — элемент точечной группы кристалла $\mathscr{F}$, называется *звездой волнового вектора*.

Покажем, что инвариантность уравнения Шредингера относительно обращения времени ($t\to -t$) приводит в некоторых случаях к дополнительной симметрии энергии $\mathscr{E}_n(\boldsymbol{k})$.

Рассмотрим временно́е уравнение Шредингера

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{\mathcal{H}}\psi \tag{9.23}$$

и возьмем ему комплексно-сопряженное

$$i\hbar \frac{\partial \psi^*}{\partial (-t)} = \hat{\mathcal{H}}\psi^*; \tag{9.24}$$

при этом мы предполагаем, что гамильтониан $\hat{\mathcal{H}}$ веществен. Мы видим, что состояние ψ^* развивается в направлении $-t$ так же, как состояние ψ в направлении t. Так как плотность вероятности пропорциональна $|\psi|^2$, то обращение времени на нее не влияет. В то же время скорости, пропорциональные $(1/i)\,\psi^*\nabla\psi$, меняют свое направление.

Для стационарного состояния

$$\hat{\mathcal{H}}\psi_n = \mathcal{E}_n\psi_n, \tag{9.24}$$

откуда

$$\hat{\mathcal{H}}\psi_n^* = \mathcal{E}_n\psi_n^*, \tag{9.24а}$$

и, следовательно, ψ_n и ψ_n^* соответствуют одной и той же энергии $\mathcal{E}_n$. Если ψ_n и ψ_n^* линейно независимы, то это приводит к дополнительному вырождению. Из (9.23) и (9.24) видно, что это дополнительное вырождение связано с симметрией по отношению к обращению времени.

Для электрона в периодическом поле волновая функция имеет вид (9.15), поэтому

$$\psi_{nkj}^*(\boldsymbol{r}) = u_{nkj}^*(\boldsymbol{r})\, e^{-i\boldsymbol{k}\boldsymbol{r}}, \tag{9.25}$$

которое отличается от состояния (9.15) заменой $\boldsymbol{k}$ на $-\boldsymbol{k}$. Так как (9.15) и (9.25) удовлетворяют одному и тому же уравнению (9.14), то

$$\mathcal{E}_n(-\boldsymbol{k}) = \mathcal{E}_n(\boldsymbol{k}). \tag{9.26}$$

Таким образом, энергетические поверхности $\mathcal{E}_n(\boldsymbol{k}) = \text{const}$ обладают центром симметрии независимо от того, обладает ли им прямая (обратная) решетка кристалла.

Если $\boldsymbol{k}$ расположен внутри зоны Бриллюэна и направлен вдоль какой-либо из ее осей или плоскостей симметрии, то среди элементов группы $\mathcal{F}$ имеются такие, что

$$R\boldsymbol{k} = \boldsymbol{k}. \tag{9.27}$$

Если же конец вектора $\boldsymbol{k}$ лежит на границе зоны Бриллюэна, то преобразование симметрии R может превратить $\boldsymbol{k}$ в эквивалентный вектор, т. е.

$$R\boldsymbol{k} = \boldsymbol{k} \pm \boldsymbol{b}_i, \tag{9.27а}$$

где $\boldsymbol{b}_i$ — основной вектор обратной решетки.

Элементы пространственной группы кристалла (9.16), у которых точечные преобразования R удовлетворяют соотношениям (9.27) или (9.27а), образуют ее подгруппу и называются *группой волнового вектора* $\boldsymbol{G_k}$. Группа $\boldsymbol{G_k}$ определяет вырождение и симметрию блоховских волновых функций в точке $\boldsymbol{k}$ бриллюэновской зоны во всех зонах (полосах) энергии n. Т. е. каждый элемент группы $\boldsymbol{G_k}$ преобразует блоховскую функцию с волновым вектором $\boldsymbol{k}$ и энергией $\mathscr{E}_n(\boldsymbol{k})$ в другие блоховские функции с тем же (или эквивалентным) волновым вектором и с той же энергией. Для применения теории групп к определению вырождения и симметрии блоховских функций в определенных точках $\boldsymbol{k}$-пространства мы должны знать в этой точке неприводимые представления группы $\boldsymbol{G_k}$ или хотя бы их характеры.

Непосредственно легко определить характеры точечной группы $\mathscr{F}_k$, элементы которой R оставляют вектор $\boldsymbol{k}$ инвариантным или преобразуют его в эквивалентный. Ниже мы установим связь между неприводимыми представлениями $\Gamma(g)$ группы волнового вектора $\boldsymbol{G_k}$ и неприводимыми представлениями $\Gamma(R)$ точечной группы $\mathscr{F}_k$. Как мы увидим дальше, для симморфных групп эти представления совпадают.

Рассмотрим, в иллюстративных целях, группу $\mathscr{F}_k$ для плоской квадратной решетки при различных положениях волнового вектора $\boldsymbol{k}$. Если сторона квадрата прямой решетки равна a, то бриллюэновская зона имеет тоже форму квадрата со стороной $2\pi/a$ (рис. II.18). Точечной группой $\mathscr{F}$ кристалла является в этом случае группа $\boldsymbol{C_{4v}}$ с восемью элементами: E, $2C_4$, C_2, $2\sigma_v$, $2\sigma_d$ (σ_v — плоскости отражения, параллельные сторонам квадрата, σ_d — плоскости отражения, проходящие через его диагонали).

В наиболее общем случае *а*) звезда волнового вектора состоит из восьми разных векторов $\boldsymbol{k}$, получающихся при применении всех восьми элементов группы $\boldsymbol{C_{4v}}$; группа $\mathscr{F}_k$ состоит при этом из одного элемента E. В случае *б*) вектор лежит в плоскости $\sigma_v(AA')$ и звезда волнового вектора состоит из четырех векторов; группа $\mathscr{F}_k$ в этом случае состоит из двух элементов: E, σ_v. В случае *в*) звезда волнового вектора тоже состоит из четырех векторов; группа $\mathscr{F}_k$ для одного из них состоит из элементов E и σ_d. В случаях *г*), *д*) и *е*) концы векторов $\boldsymbol{k}$ в звезде расположены на границе зоны Бриллюэна. Так как вектор k и $k' = k + \frac{2\pi}{a}$ эквивалентны, то в случае *г*) $\boldsymbol{k}$-звезда состоит из четырех векторов, а группа $\mathscr{F}_k$ из двух элементов: E и σ_v. В случае *д*) звезда волнового вектора состоит из двух векторов, а группа $\mathscr{F}_k$ из четырех элементов: E, C_2, $2\sigma_v$. Наконец, в случае *е*) все четыре изображенных вол-

новых вектора эквивалентны, а группа $\mathcal{F}$ совпадает с точечной группой кристалла C_{4v}.

4. Покажем, что если $\Gamma(R)$ — неприводимое представление точечной группы $\mathcal{F}_k$, то неприводимое представление группы волнового вектора G_k равно (ограничение смотри ниже)

$$\Gamma(g)=\Gamma(\{R\,|\,\alpha(R)+a\})=\Gamma(R)\exp i\,[k(\alpha+a)]. \quad (9.28)$$

Здесь $g=\{R\,|\,\alpha(R)+a\}$ — элемент группы G_k, $a\equiv a_n$ — вектор решетки и $\alpha(R)\equiv\alpha$ — несобственная трансляция, соответствующая элементу R.

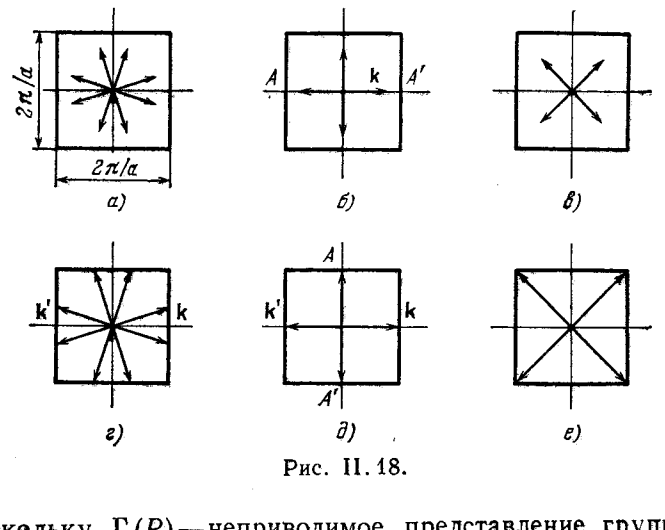


Рис. II.18.

Поскольку $\Gamma(R)$ — неприводимое представление группы $\mathcal{F}_k$:

$$\Gamma(R_2R_1)=\Gamma(R_2)\,\Gamma(R_1). \quad (9.29)$$

Из (9.28) следует, что

$$\Gamma(g_2)\,\Gamma(g_1)=\Gamma(R_2)\,\Gamma(R_1)\exp i\,[k(\alpha_2+a_2)+k(\alpha_1+a_1)]. \quad (9.30)$$

Здесь a_1 и $\alpha_1\equiv\alpha(R_1)$ — вектор решетки и несобственная трансляция, соответствующие элементу R_i для элемента $g_1\in G_k$; аналогичный смысл имеют величины a_2 и α_2.

Для элемента g_2g_1 группы G_k из (9.28) следует

$$\begin{aligned}\Gamma(g_2g_1)&=\Gamma(\{R_2\,|\,\alpha_2+a_2\}\{R_1\,|\,\alpha_1+a_1\})=\\&=\Gamma(\{R_2R_1\,|\,R_2\alpha_1+R_2a_1+\alpha_2+a_2\})=\\&=\Gamma(R_2R_1)\exp i\,[k(R_2\alpha_1+R_2a_1+\alpha_2+a_2)],\end{aligned} \quad (9.31)$$

где мы для произведения g_2g_1 воспользовались выражением (5.8). Заменяя здесь, согласно (9.29), $\Gamma(R_2R_1)$ на произведение $\Gamma(R_2)\,\Gamma(R_1)$ и выражая каждый из этих множителей посредством (9.28),

получим

$$\Gamma(g_2g_1)=\Gamma(g_2)\,\Gamma(g_1)\exp i[\boldsymbol{k}(R_2\boldsymbol{a}_1+R_2\boldsymbol{a}_1)-\boldsymbol{k}(\boldsymbol{a}_1+\boldsymbol{a}_1)].$$

Используя для первых двух слагаемых показателя экспоненты преобразование (9.20), получим

$$\Gamma(g_2g_1)=\Gamma(g_2)\,\Gamma(g_1)\exp[i(R_2^{-1}\boldsymbol{k}-\boldsymbol{k})(\boldsymbol{a}_1+\boldsymbol{a}_1)]. \tag{9.32}$$

Если волновой вектор лежит внутри зоны Бриллюэна, то

$$R_2^{-1}\boldsymbol{k}=\boldsymbol{k}. \tag{9.33}$$

В этом случае экспонента в (9.32) равна единице и, следовательно,

$$\Gamma(g_2g_1)=\Gamma(g_2)\,\Gamma(g_1), \tag{9.34}$$

т. е. $\Gamma(g)$, определяемое равенством (9.28), — неприводимое представление группы волнового вектора $G_{\boldsymbol{k}}$.

Если вектор $\boldsymbol{k}$ кончается в некоторой точке поверхности зоны Бриллюэна, то для некоторых R_2^{-1} эквивалентный ему волновой вектор

$$R_2^{-1}\boldsymbol{k}=\boldsymbol{k}+\boldsymbol{b}, \tag{9.35}$$

где $\boldsymbol{b}$ — вектор обратной решетки. Для симморфной пространственной группы $\boldsymbol{\alpha}_1=\boldsymbol{\alpha}(R_1)=0$ и экспонента в (9.32) тоже равна единице; в самом деле,

$$e^{i\left[(R_2^{-1}\boldsymbol{k}-\boldsymbol{k})\,\boldsymbol{a}_1\right]}=e^{ib\boldsymbol{a}_1}=e^{2\pi i\times\text{целое число}}=1; \tag{9.36}$$

таким образом, и в этом случае $\Gamma(g)$ (9.29) неприводимое представление группы волнового вектора $G_{\boldsymbol{k}}$. Сложнее случай, когда конец вектора $\boldsymbol{k}$ расположен на границе зоны Бриллюэна, но пространственная группа не симморфна. Этот случай мы здесь рассматривать не будем (см. гл. IV, § 9).

§ 10. Правила отбора

Вероятность перехода системы из состояния с волновой функцией $\psi_{i\alpha}$ (α-я волновая (базисная) функция i-го уровня энергии, Γ_i неприводимого представления) в состояние с волновой функцией $\psi_{k\beta}$ под действием возмущения $\hat{Q}$ пропорциональна квадрату модуля матричного элемента

$$M=\int\psi^*_{k\beta}\hat{Q}\psi_{i\alpha}\,d\tau, \tag{10.1}$$

где $d\tau$ — произведение дифференциалов координат конфигурационного пространства системы.

Во многих случаях нам неизвестны волновые функции системы, так что мы не можем вычислить матричный элемент (10.1). Однако часто нам достаточно знать, равен ли матричный эле-

мент нулю или отличен от нуля. Покажем, как теория групп позволяет ответить на этот вопрос, т. е. позволяет сформулировать *правила отбора*.

Для простого случая оптического электрона в свободном атоме, когда возмущением является электрическое поле световой волны, такие правила отбора известны из квантовой механики. Для дипольных переходов изменение азимутального квантового числа l должно равняться ± 1 (т. е. $\Delta l = \pm 1$), а магнитного квантового числа ± 1 или 0 (т. е. $\Delta m = \pm 1$ или $\Delta m = 0$) [1]. Для рассмотрения более общего случая покажем, что

$$\int \psi_{i\alpha}\, d\tau = 0, \tag{10.2}$$

если неприводимое представление Γ_i не является единичным.

Подвергая подынтегральную функцию в (10.2) ортогональному преобразованию симметрии $\hat{P}_R$, где R — элемент группы уравнения Шредингера, получим

$$\int \psi_{i\alpha}\, d\tau = \int \hat{P}_R \psi_{i\alpha}\, d\tau = \int \sum_{\beta} \Gamma_i (R)_{\beta\alpha} \psi_{i\beta}\, d\tau. \tag{10.3}$$

Здесь первое равенство основано на том, что якобиан ортогонального преобразования $\boldsymbol{x}' = \boldsymbol{Rx}$ равен единице, а второе — на (7.15).

Суммируя обе части (10.3) по всем элементам группы, получим

$$h \int \psi_{i\alpha}\, d\tau = \sum_{\beta} \int \psi_{i\beta}\, d\tau \cdot \sum_{R} \Gamma_i (R)_{\beta\alpha}, \tag{10.4}$$

где h — порядок группы.

Если Γ_i не есть единичное представление, то

$$\sum_{R} \Gamma_i (R)_{\beta\alpha} = 0.$$

Это следует из (6.19), если предположить, что представление j есть единичное, а i — нет. Таким образом, правая часть равенства (10.4) равна нулю, следовательно, (10.2) доказано.

Пусть отдельные множители (функции) подынтегрального выражения матричного элемента (10.1) преобразуются по неприводимым представлениям Γ_k, Γ_Q и Γ_i группы гамильтониана. Тогда произведение этих множителей есть базисная функция приводимого, в общем случае, представления прямого произведения $\Gamma_k \times \Gamma_Q \times \Gamma_i$ (см. § 7, п. 3). Если это прямое произведение не содержит единичного представления, то его базисные функции

---

[1]) Блохинцев Д. И. Основы квантовой механики. — 5 изд. — М.: Наука, 1976, § 90.

не содержит базисной функции единичного представления, а тогда в соответствии с (10.2) матричный элемент (10.1) равен нулю.

Таким образом, для определения правил отбора достаточно определить, содержится ли в прямом произведении $\Gamma_k \times \Gamma_Q \times \Gamma_i$ единичное представление. При этом число единичных представлений, содержащихся в этом произведении, определяет число линейно независимых матричных элементов. Этот рецепт может быть упрощен на основании следующей теоремы.

Прямое произведение двух различных неприводимых представлений $\Gamma_l \times \Gamma_m$ не содержит единичного представления, а прямое произведение двух одинаковых неприводимых представлений $\Gamma_l \times \Gamma_l$ содержит единичное представление один раз. Приведем ее доказательство.

Для определения того, сколько раз единичное представление содержится в приводимом представлении, воспользуемся формулой (6.34); полагая в ней $a_j = a_1$ и $\chi_j(R)^* = 1$ (характеры единичного представления для всех R равны единице), получим

$$a_1 = \frac{1}{h} \sum_R \chi(R).$$

Для прямого произведения представлений $\Gamma_l \times \Gamma_m$, согласно (7.18), характер

$$\chi(R) = \chi_l(R)\,\chi_m(R),$$

поэтому

$$a_1 = \frac{1}{h} \sum_R \chi_l(R)\,\chi_m(R).$$

Отсюда по теореме об ортогональности характеров (6.26) следует, что

$$a_1 = 0, \text{ если } l \neq m,$$
$$a_1 = 1, \text{ если } l = m.$$

Таким образом, если при разложении на неприводимые представления Γ_k и $\Gamma_Q \times \Gamma_i$ они содержат общие неприводимые представления, то матричный элемент M (10.1) отличен от нуля, в противоположном случае он равен нулю.

2. Рассмотрим в качестве примера правила отбора для дипольных переходов (в этом случае $\hat{Q}$ = радиусу-вектору $= \boldsymbol{r} = \{x, y, z\}$) в поле кубической симметрии $\boldsymbol{O}$.

Как будет показано в гл. IV § 8, радиус-вектор $\boldsymbol{r}$, т. е. координаты x, y и z преобразуются по неприводимому представлению Γ_{15} или F_1 группы $\boldsymbol{O}$.

Пользуясь формулой (7.18), составим таблицу характеров (табл. 11.10) прямого произведения $F_1 \times \Gamma_i$, где Γ_i — одно из неприводимых представлений группы $\boldsymbol{O}$, соответствующее базисным функциям $\psi_{i\alpha}$.

Таблица II.10

$F_1\times\Gamma_i$	E	$8\,(C_3;\ C_3^2)$	$3C_4^2$	$6C_2$	$6C_4$
$F_1\times A_1$	3	0	—1	—1	1
$F_1\times A_2$	3	0	—1	1	—1
$F_1\times E$	6	0	—2	0	0
$F_1\times F_1$	9	0	1	1	1
$F_1\times F_2$	9	0	1	—1	—1

Разложим прямые произведения $F_1\times\{A_1,\ A_2,\ E,\ F_1,\ F_2\}$ на неприводимые представления группы $\boldsymbol{O}$; пользуясь формулой (6.34), получим

$$\begin{gathered}F_1\times A_1=F_1,\quad F_1\times A_2=F_2,\quad F_1\times E=F_1+F_2,\\ F_1\times F_1=A_1+E+F_1+F_2,\quad F_1\times F_2=A_2+E+F_1+F_2.\end{gathered}\tag{10.5}$$

Пользуясь доказанными выше теоремами, мы видим, что матричные элементы M отличны от нуля, т. е. разрешены дипольные переходы между состояниями, соответствующими второму представлению в прямом произведении в левой части равенства, и состояниями, неприводимыми представлениями в правой части равенства. То есть разрешены переходы между состояниями

$$F_1\leftrightarrow A_1,\ E,\ F_1,\ F_2;\qquad F_2\leftrightarrow A_2,\ E,\ F_1,\ F_2,$$

что непосредственно следует из двух последних равенств (10.5) (первые три равенства ничего нового не добавляют). В то же время запрещены переходы между $F_1\rightleftarrows A_2$, $F_2\rightleftarrows A_1$, $A_1\rightleftarrows A_2$ и $E\rightleftarrows A_2$.

ГЛАВА III

КОЛЕБАНИЯ АТОМОВ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ

§ 1. Природа сил взаимодействия атомов в кристалле

1. То обстоятельство, что в определенных условиях атомы могут образовывать устойчивые молекулы и кристаллы, свидетельствует о том, что между ними могут существовать силы притяжения, которые уравновешиваются на расстояниях порядка 10^{-8} *см* силами отталкивания. Почти во всех случаях представляется более удобным оперировать не с силами, а с потенциальной энергией взаимодействия атомов $\mathcal{U}(R)$, которую мы предполагаем зависящей только от расстояния R между ядрами атомов. Кривые *1* и *2* на рис. III.1 изображают возможные случаи взаимодействия двух атомов, один из которых помещен в начале координат 0, а другой A может перемещаться вдоль оси R. Так как потенциальная энергия определяется с точностью до постоянного слагаемого, то всегда можно положить $\mathcal{U}=0$ при $R\to\infty$. Сила, действующая на атом A, равна $\boldsymbol{F}=-\operatorname{grad}\mathcal{U}(R)=-\frac{d\mathcal{U}}{dR}\frac{\boldsymbol{R}}{R}$, где $\boldsymbol{R}=\overrightarrow{0A}$ — радиус-вектор, проведенный из 0 в A. Таким образом, в тех точках, где $\frac{d\mathcal{U}}{dR}>0$ ($\boldsymbol{F}$ антипараллельно $\boldsymbol{R}$), имеют место силы притяжения, а в тех, где $\frac{d\mathcal{U}}{dR}<0$ ($\boldsymbol{F}$ параллельно $\boldsymbol{R}$), — силы отталкивания. Из рис. III.1 видно, что кривая *1* соответствует случаю, когда атомы на любом расстоянии R отталкиваются друг от друга. Кривая *2* соответствует более сложному случаю, когда при $R>R_0$ атомы притягиваются, а при $R<R_0$ отталкиваются друг от друга. $\left(\frac{d\mathcal{U}}{dR}\right)_{R_0}=0$ в точке $R=R_0$, т. е. сила $\boldsymbol{F}=-\left(\frac{d\mathcal{U}}{dR}\right)_{R_0}\frac{\boldsymbol{R}}{R}=0$ и система из

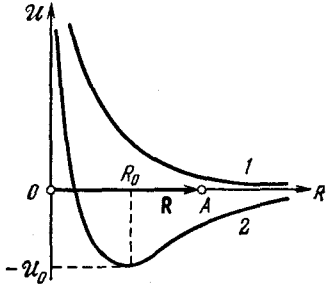


Рис. III.1.

обоих атомов находится в состоянии устойчивого равновесия. В этом случае, как это непосредственно видно из формы кривой 2, $\left(\frac{d^2\mathcal{U}}{dR^2}\right)_{R_0} \equiv \beta > 0$. При малых отклонениях атома A от положения равновесия

$$\mathcal{U}(R) = \mathcal{U}(R_0) + \left(\frac{d\mathcal{U}}{dR}\right)_{R_0}(R-R_0) + \frac{1}{2}\left(\frac{d^2\mathcal{U}}{dR^2}\right)_{R_0}(R-R_0)^2 + \\ + \frac{1}{6}\left(\frac{d^3\mathcal{U}}{dR^3}\right)_{R_0}(R-R_0)^3 + \ldots \quad (1.1)$$

с точностью до величин третьего порядка малости. Если, как это имеет место на самом деле, силы отталкивания вблизи точки $R = R_0$ возрастают быстрее, чем убывают силы притяжения, то $\left(\frac{d^3\mathcal{U}}{dR^3}\right)_{R_0} \equiv -2\gamma < 0$. Обозначая $\mathcal{U}(R_0) \equiv -\mathcal{U}_0$ и отклонение атома A от положения равновесия $R - R_0 \equiv x$, получим

$$\mathcal{U}(x) + \mathcal{U}_0 = {}^1/_2\beta x^2 - {}^1/_3\gamma x^3. \quad (1.2)$$

Сила, действующая на атом A при движении его вблизи положения равновесия вдоль оси R, равна

$$F = -\frac{d\mathcal{U}}{dx} = -\beta x + \gamma x^2. \quad (1.3)$$

Заметим, что сила, рассматриваемая в приближении $F = -\beta x$, называется квазиупругой. Можно думать, что по порядку величины $\beta \approx \gamma R_0$.

2. Последовательная теория взаимодействия атомов (ионов) должна основываться на квантовомеханическом рассмотрении движения их электронов. Атомные ядра можно при этом считать неподвижными в силу их большой массы (адиабатическое приближение). Полная энергия электронов зависит при этом от положения ядер как от параметров. При изменении расстояния между атомами полная энергия электронов, наряду с кулоновским отталкиванием ядер, играет роль потенциальной энергии взаимодействия атомов. В некоторых случаях можно приближенно понять природу взаимодействия атомов на основании более элементарных соображений, связанных со статистическим рассмотрением их электронов.

Рассмотрим с этой точки зрения природу сил отталкивания между атомами (ионами). Если исключить тривиальный случай кулоновского взаимодействия между одноименными ионами, проявляющегося на больших расстояниях между ними, то отталкивание между атомами (ионами) на близких расстояниях возникает при взаимном проникновении их электронных оболочек. Это отталкивание связано в основном с увеличением кинетической энергии атомных электронов в силу принципа Паули.

Для выяснения природы этих сил будем рассматривать электроны атомов (ионов) как вырожденный ферми-газ при абсолютном нуле температуры. Простой расчет показывает (Приложение 4), что плотность кинетической энергии электронов (кинетическая энергия в расчете на 1 $см^3$) равна $\mathscr{E}_k = \frac{3^{5/3}\pi^{4/3}\hbar^2}{10m} n^{5/3}$, где n — концентрация электронов (в данной точке), $2\pi\hbar$ — постоянная Планка и m — масса электрона. Будем в нулевом приближении считать, что плотности электронов n_a и n_b свободных атомов (ионов) a и b (рис. III.2) не меняются при взаимном проникновении электронных оболочек. Это предположение соответствует при квантовомеханических расчетах использованию невозмущенных волновых функций. В этом случае изменение плотности кинетической энергии в области наложения электронных оболочек (заштрихованная область на рис. III.2) равно

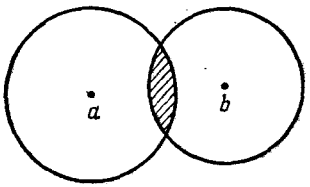


Рис. III.2.

$$\Delta\mathscr{E}_k = \frac{3^{5/3}\pi^{4/3}\hbar^2}{10m}\left[(n_a + n_b)^{5/3} - n_a^{5/3} - n_b^{5/3}\right]. \qquad (1.4)$$

Легко видеть, что эффект наложения электронных оболочек атомов a и b связан с увеличением кинетической энергии электронов в системе, т. е. с возникновением сил отталкивания (тем бо́льших, чем меньше расстояние между a и b). К ним следует добавить чисто кулоновские силы отталкивания между ядрами атомов a и b, возникающие при проникновении ядра одного атома в электронную оболочку другого. Более подробный анализ показывает, что учет обменных эффектов в рамках статистической теории приводит при взаимном проникновении электронных оболочек к появлению некоторых сил притяжения, не меняющих, однако, качественной картины, описанной выше. Квантовомеханическая теория сил отталкивания атомов дает для потенциальной энергии выражение вида $A\exp(-R/a)$, где A и a — положительные константы.

3. Для образования устойчивых кристаллов, наряду с силами отталкивания между атомами (ионами), должны существовать силы притяжения между ними. Обычно рассматриваются четыре основных типа связей в кристаллах: а) *ионные* или *гетерополярные*, б) *ковалентные* или *гомеополярные*, в) *ван-дер-ваальсовские* или *дисперсионные* и г) *металлические*.

Следует отметить, что в большинстве случаев связи в кристаллах носят смешанный характер, поэтому нередки утверждения, что связь, например, на столько-то процентов ковалентна, а на столько-то гетерополярна. Когда мы говорим о связи одного вида, это значит, что этот тип связи превалирует.

В наиболее чистом виде гетерополярная связь осуществляется в ионных кристаллах, например щелочно-галоидных соединениях NaCl, KI, CsCl. При этом в первом приближении взаимодействие ионов рассматривается как взаимодействие точечных зарядов, помещенных в узлах решетки. Так как ионы первой координационной группы имеют всегда другой знак, чем центральный ион, то кулоновское взаимодействие всех ионов решетки приводит к некоторому результирующему притяжению, обеспечивающему стабильность решетки.

Ввиду медленности убывания кулоновских сил с расстоянием в количественной теории необходимо учитывать взаимодействие центрального иона и с более далекими ионами обоих знаков. В следующем приближении учитывается взаимная поляризация ионов.

В гл. I, § 2, п. 4 мы рассмотрели решетки некоторых ионных кристаллов. В них число положительных ионов равно числу отрицательных, поэтому как кристалл в целом, так и элементарная кристаллическая ячейка нейтральны. При приложении механического напряжения к ионным кристаллам они ведут себя по-разному.

В некоторых кристаллах, называемых *пьезоэлектриками*, вектор поляризации P_i $(i=x, y, z)$ линейно зависит от компонент тензора напряжения σ_{kl} [1]), т. е.

$$P_i=\sum_{k,l}\gamma_{ikl}\sigma_{kl}. \tag{1.5}$$

Здесь γ_{ikl} — пьезоэлектрический тензор 3-го ранга (см. Приложение 2) (*пьезоэлектрический модуль*).

Как будет показано ниже, в кристаллах с центром инверсии (например, NaCl, CsBr) $\gamma_{ikl}=0$, поэтому в таких кристаллах пьезоэлектрический эффект равен нулю. Если к пьезоэлектрическому кристаллу приложить электрическое поле E_i, в нем возникает деформация, пропорциональная первой степени электрического поля, т. е. компоненты тензора деформации [2])

$$u_{kl}=\sum\gamma_{ikl}E_i. \tag{1.5а}$$

Используя термодинамику, можно показать, что тензоры γ_{ikl} в (1.5) и (1.5а) совпадают [3]). Деформация кристалла в электрическом поле (1.5а) получила название *обратного пьезоэлектрического эффекта*. Так как тензор деформации u_{kl} симметричен, то пьезоэлектрический тензор γ_{ikl} симметричен по индексам k и l,

---

[1]) Ландау Л. Д., Лифшиц Е. М./Теоретическая физика, т. 7.—Теория упругости, —М.: Наука, 1965, § 2.

[2]) Ландау Л. Д., Лифшиц Е. М. Там же § 1.

[3]) Най Дж. Физические свойства кристаллов —М.: ИЛ, 1960, гл. X.

т. е. $\gamma_{ikl}=\gamma_{ilk}$ Из-за последнего условия максимальное число независимых компонент тензора γ_{ikl} равно $3^3-3^2=27-9=18$.

Покажем, что для кристаллов с центром инверсии все компоненты тензора γ_{ikl} равны нулю, т. е. такие кристаллы непьезоэлектрики.

Компоненты тензора γ_{ikl} преобразуются как произведение прямоугольных координат $x_ix_kx_l$. При преобразовании инверсии новые («штрихованные») координаты равны

$$x_1'=-x_1,\quad x_2'=-x_2,\quad x_3'=-x_3, \tag{1.6}$$

т. е. направляющие косинусы $\alpha_{ik}=-\delta_{ik}$, где δ_{ik}—символ Кронекера. Таким образом, при преобразовании инверсии

$$\gamma_{ikl}'=\sum_{mnp}\alpha_{im}\alpha_{kn}\alpha_{lp}\gamma_{mnp}=-\sum_{mnp}\delta_{im}\delta_{kn}\delta_{lp}\gamma_{mnp}=-\gamma_{ikl}. \tag{1.6а}$$

Но преобразование инверсии является преобразованием симметрии кристалла, поэтому материальный тензор γ_{ikl} не может измениться, т. е. $\gamma_{ikl}'=\gamma_{ikl}$; отсюда и из (1.6а) следует, что $\gamma_{ikl}=0$.

Рассмотрим теперь пьезоэлектрические свойства кристалла цинковой обманки ZnS, кристаллизующейся в решетке типа германия (рис. I.13), не обладающей центром инверсии. Точечная группа симметрии кристалла ZnS—группа тетраэдра T_d. Используя табл. IV.2, дающую преобразование координат $x=x_1$, $y=x_2$, $z=x_3$ при действии элементов группы T_d [1]), применим преобразование $R_2(x_1, \bar{x}_2, \bar{x}_3)$ к компоненте тензора γ_{112}, которая преобразуется как $x_1^2x_2$. Имеем $\gamma_{112}'=-\gamma_{112}$ и $\gamma_{112}'=\gamma_{112}$, откуда $\gamma_{112}=0$. Применяя то же преобразование $R_2(x_1, \bar{x}_2, \bar{x}_3)$, можно показать, что и

$$\gamma_{211}=\gamma_{113}=\gamma_{311}=\gamma_{222}=\gamma_{333}=0.$$

Используя преобразование $R_3(\bar{x}_1, x_2, \bar{x}_3)$, можно показать, что

$$\gamma_{122}=\gamma_{212}=\gamma_{223}=\gamma_{322}=\gamma_{111}=0,$$

а из $R_4(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ следует, что

$$\gamma_{133}=\gamma_{331}=\gamma_{233}=\gamma_{323}=0.$$

Таким образом, из восемнадцати независимых компонент тензора γ_{ikl} отличны от нуля только три: γ_{123}, γ_{213} и γ_{312}. Используя преобразования $R_9(x_3, x_1, x_2)$ и $R_{15}(\bar{x}_3, \bar{x}_2, \bar{x}_1)$, можно показать, что

$$\gamma_{123}=\gamma_{213}=\gamma_{312}\equiv\gamma_0. \tag{1.7}$$

Таким образом, не равные нулю компоненты тензора γ_{ikl} равны между собой. Мы видим, что пьезоэлектрические свойства кристалла ZnS характеризуются одной материальной константой γ_0.

[1]) Смысл символов $R_i(\ldots)$ объяснен в гл. IV, § 8, п. 2.

4. Ковалентная связь возникает между нейтральными атомами на близких расстояниях ($\sim 10^{-8}$ *см*) при выполнении определенных условий. В наиболее простом виде ковалентная связь реализуется между атомами водорода в молекуле H_2 (В. Гайтлер и Ф. Лондон, 1927). Ковалентная связь не может быть интерпретирована на языке понятий классической физики. Для объяснения ковалентной связи существенны квантовомеханические особенности поведения системы одинаковых частиц (электронов)[1]. Классическая физика была совершенно бессильна в объяснении свойства насыщаемости ковалентной связи, т. е. неспособности, например, атома водорода к присоединению более чем одного другого атома водорода. Это свойство характеризуется в химии понятием *валентности*; оно связано со спариванием электронов обоих атомов и образованием синглетного состояния, в котором спины электронов антипараллельны (для триплетного состояния, в котором спины электронов параллельны, кривая взаимодействия атомов имеет вид *1* на рис. III.1, т. е. атомы на всех расстояниях отталкиваются друг от друга).

Ковалентная связь может осуществляться не только между двумя атомами водорода, но и между другими атомами, обладающими электронами, способными к образованию пар с антипараллельными спинами.

Например, атом азота N имеет два электрона в 1s-, два электрона в 2s- и три электрона в 2p-состоянии, т. е. имеет электронную конфигурацию, которую обозначают $(1s)^2\,(2s)^2\,(2p)^3$. Спектроскопические данные показывают, что спины трех электронов в 2p-состоянии параллельны, т. е. спиново не насыщены, следовательно, они способны образовать три ковалентные связи; таким образом, азот трехвалентен.

Опыт подтверждает это. Так, например, при соединении азота с водородом получается аммиак NH_3. Аналогично образуется двухатомная молекула азота N_2, в которой атомы связаны тремя парами электронов с антипараллельными спинами. Три электрона в атоме азота в 2p-состоянии имеют магнитные квантовые числа $m = +1, -1, 0$ и им соответствуют волновые функции вида $\psi_{+1} = xf(r)$, $\psi_{-1} = yf(r)$ и $\psi_0 = zf(r)$, т. е. электронные облака трех валентных электронов вытянуты по трем взаимно перпендикулярным направлениям x, y и z[2]. Энергетический выигрыш при образовании ковалентной связи существенно связан с перекрытием волновых функций электронов соответствующей пары с антипараллельными спинами. Таким образом, следует думать, что атомы водорода в молекуле аммиака будут распо-

[1]) Блохинцев Д. И. Основы квантовой механики.—М.: Наука, 1976, § 125.

[2]) Гайтлер В. Элементарная квантовая механика.—М.: Гостехиздат, 1948, с. 36.

ложены по отношению к атому азота по трем взаимно перпендикулярным направлениям (*направленные валентности*). Опыт показывает, что молекула NH_3 действительно имеет форму пирамиды с углом $\widehat{HNH}$, близким к 90° и равным 109°. Несколько бо́льший угол между направленными валентностями в молекуле NH_3 объясняется взаимным отталкиванием атомов водорода.

Атом углерода C имеет электронную конфигурацию $(1s)^2(2s)^2(2p)^2$. Так как в s-состояниях спины насыщены (антипараллельны), то атом углерода должен был бы быть двухвалентным. Такой вывод находится, однако, в противоречии с данными всей органической химии, согласно которой атом углерода четырехвалентен. Более тщательный теоретический и экспериментальный анализ вопроса показывает, что атом углерода участвует в химических соединениях не в основном, а в возбужденном состоянии: $(1s)^2(2s)^1(2p)^3$. В этом случае спины всех четырех электронов в 2s- и 2p-состояниях не насыщены (параллельны) и могут участвовать в образовании ковалентной связи. Если говорить точнее, то атом углерода образует ковалентную связь в электронном состоянии, являющемся с у п е р п о з и ц и е й одного 2s- и трех 2p-состояний. Коэффициенты (веса) каждого из этих состояний в линейной суперпозиции и направления четырех валентных связей определяются из условия минимума энергии образования молекулы. Математический анализ, который мы не имеем возможности здесь привести, показывает, что направления валентных связей совпадают с направлениями *O1*, *O2*, *O3* и *O4* в тетраэдре (рис. I.11). При этом, как мы знаем, углы между этими направлениями равны 109°,5. Опыт показывает, что молекула метана CH_4 действительно имеет указанную выше тетраэдрическую структуру. Направленная четырехвалентность атомов углерода проявляется при образовании кристалла алмаза, в котором каждый атом C расположен в центре тетраэдра, образованного четырьмя другими C-атомами (рис. I.13). Атом кремния Si имеет 4 электрона в *M*-оболочке в состоянии $(3s)^2(3p)^2$, и так как *K*- и *L*-оболочки спиново насыщены, то он будет вести себя подобно атому углерода. Аналогичными свойствами обладают атом германия Ge, имеющий 4 электрона в *N*-оболочке в состоянии $(4s)^2(4p)^2$, атом олова Sn, имеющий 4 электрона в *O*-оболочке в состоянии $(5s)^2(5p)^2$, и атом свинца Pb, имеющий 4 электрона в *P*-оболочке в состоянии $(6s)^2(6p)^2$. Действительно, кремний, германий и серое олово также кристаллизуются в решетке типа алмаза и принадлежат к типичным валентным или атомным кристаллам. Что же касается обычного (не серого) олова и свинца, то в них ковалентный характер связи атомов не проявляется, будучи подавленным «металличностью» соединения (см. ниже).

Экспериментальные исследования последних лет показали, что химические соединения типа $A^{III}B^{V}$, т. е. элементов III и V групп периодической системы, обладают многими свойствами (кристаллическая решетка, структура электронных зон), подобными элементами IV группы — германию и кремнию. К числу таких соединений относятся InSb и GaAs. Индий имеет в O-оболочке три электрона в конфигурации $(5s)^2(5p)^1$, а сурьма — пять электронов в состоянии $(5s)^2(5p)^3$. Таким образов, как и в случае двух атомов Ge или Si имеется 4 электрона в s-состоянии и 4 — в p-состоянии. Если один p-электрон Sb частично перейдет к In, то может возникнуть ковалентная связь, подобная той, которая имеет место в кристаллах Si и Ge. Аналогичное положение имеет место в соединении GaAs. Конечно, для InSb и GaAs связь является не чисто гомеополярной, а частично носит ионный характер.

5. Ковалентная связь образуется только при выполнении определенных условий. Во-первых, необходимо, чтобы у атомов имелись такие электроны, которые могли бы образовать при сближении атомов пары с противоположно направленными спинами (синглетное состояние). Во-вторых, расстояние между атомами должно быть так мало, чтобы проявились квантомеханические свойства, связанные с принципиальной неразличимостью тождественных частиц, входящих в одну систему. Расчет показывает, что ковалентные силы быстро (экспоненциально) убывают с расстоянием. На этих расстояниях между любыми атомными системами начинают проявляться некоторые универсальные силы притяжения. Эти силы называются *ван-дер-ваальсовскими*, или *дисперсионными*, так как, с одной стороны, они являются причиной отступления в поведении реальных газов от идеальных, а с другой — входящие в них параметры определяют дисперсию света атомами. В тех случаях, когда по определенным причинам (замкнутость электронных оболочек, насыщенность валентных связей в молекулах, взаимодействие которых рассматривается и т. п.) отсутствуют ковалентные, ионные и металлические (см. ниже) связи, ван-дер-ваальсовское взаимодействие является причиной связи частиц в твердых телах (кристаллы аргона, криптона, ксенона, молекулярные кристаллы).

Если расстояние между атомными системами $R \gg a$ (a — размер атомной системы), то ван-дер-ваальсовские (дисперсионные) силы могут быть вычислены во втором приближении квантовомеханической теории возмущений. Можно показать[1]), что ван-дер-ваальсовская энергия взаимодействия двух атомных систем

$$\mathcal{U}(R) = -W_0/R^6. \tag{1.8}$$

[1]) Блохинцев Д. И., § 127.

Здесь R — расстояние между системами, а

$$W_0 \approx \frac{3}{2} \frac{J_a J_b}{J_a + J_b} \alpha_a \alpha_b, \tag{1.9}$$

где J_a, J_b — потенциалы ионизации, а α_a, α_b — поляризации атомов (молекул) a и b.

Ван-дер-ваальсовские силы, так же как ионные, не обладают насыщаемостью, характерной для ковалентной связи.

Наряду с рассмотренными выше ионными, ковалентными и ван-дер-ваальсовскими силами существуют *дипольными* и *индукционные* силы, обязанные наличию постоянного электрического дипольного момента у молекул. Они могут иметь значение для сложных молекулярных решеток, и мы рассматривать их не будем.

В кристаллах, в которых атомы или молекулы связаны ван-дер-ваальсовскими силами (инертные газы при низких температурах, молекулярные кристаллы), условие $R \gg a$ не выполняется, поэтому формула (1.6) в лучшем случае может служить для качественных оценок.

6. Типичные металлы, такие, например, как Li, Na, K, Cu, Ag, Fe, Ni и другие, обладают рядом характерных электрических, оптических и механических свойств. Все они характеризуются относительно большой электропроводностью и коэффициентом поглощения света и высокой пластичностью и ковкостью. Эти свойства однозначно указывают на то, что в металлах имеется большое количество (порядка числа атомов) так называемых «свободных» электронов, т. е. таких электронов, которые могут и в слабых электрических внешних полях перемещаться по объему кристалла на макроскопические расстояния. Наиболее простая модель металла, предложенная Друде, представляет собой совокупность положительно заряженных ионов, расположенных в узлах кристаллической решетки и идеального газа «свободных» электронов, движущихся между ионами. Несмотря на все изменения и усложнения современной электронной теории металлов (которых мы частично коснемся ниже), такая схематическая модель не потеряла своего значения, если мы только учтем, что идеальный газ «свободных» электронов при всех практически достижимых температурах металла сильно вырожден.

Как было показано впервые Я. И. Френкелем (дальнейшие усовершенствования его идеи не представляются нам существенными), силы сцепления в металлическом кристалле могут быть объяснены на основе той же простой модели «свободных» электронов [1]). Френкель рассматривает полную энергию связи металла,

[1]) Френкель Я. И. Введение в теорию металлов.—М., 1958, гл. II.

состоящей из двух частей: отрицательной энергии кулоновского взаимодействия «свободных» электронов с положительными ионами (она $\sim 1/a$, где a — постоянная решетки) и положительной кинетической энергии вырожденного газа «свободных» электронов (она $\sim n^{2/3} \sim 1/a^2$, где n — концентрация «свободных» электронов, см. Приложение 4).

Очевидно, что полная энергия связи имеет минимальное значение при некотором значении a, которое, как показывает простой расчет, по порядку величины равно 10^{-8} *см*.

§ 2. Колебания и волны в простой одномерной (линейной) решетке

1. Атомы кристалла покоятся в узлах решетки только при абсолютном нуле температуры [1]). При повышении температуры атомы начинают колебаться около своих положений устойчивого равновесия, поэтому нам необходимо рассмотреть динамику движения атомов в кристалле.

Мы начнем изучение вопроса с рассмотрения простейшего случая колебания одинаковых атомов в одномерной (линейной)

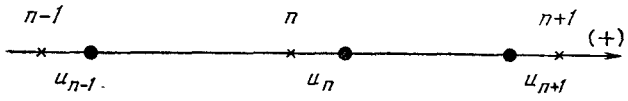


Рис. III.3.

решетке, подчиняющихся в своем движении законам классической механики. Как мы увидим впоследствии, почти все закономерности, полученные для такой схематической одномерной модели, оправдываются и для трехмерных решеток. Кроме того, выяснится, что при достаточно высоких температурах движение атомов в кристалле действительно подчиняется законам классической механики.

На рис. III.3 изображена линейная цепочка из одинаковых атомов с массами m, отклоненных от равновесных узлов с номерами $(n-1)$, n, $(n+1)$ на величины $u_{n-1} > 0$, $u_n > 0$, $u_{n+1} < 0$. В одномерном случае будем учитывать взаимодействие только ближайших (соседних) атомов, что не отражается существенным образом на результатах. Отклонения u_n и силы, действующие на атомы, считаются положительными, если их направления совпадают с направлением положительной оси и отрицательными — в противоположном случае.

---

[1]) Мы отвлекаемся здесь от наличия квантовых эффектов, на самом деле существенных при низких температурах.

Как было показано в гл. III, § 1, п. 1, при малых отклонениях атомов от положений равновесия ($|u_n| \ll a$ — здесь расстояние между узлами), силы взаимодействия можно рассматривать как квазиупругие, т. е. пропорциональные изменению расстояния между атомами. Таким образом, силы, действующие на n-й атом со стороны $(n-1)$-го и $(n+1)$-го атомов, равны: $f_{n,\,n-1} = -\beta(u_n - u_{n-1})$ и $f_{n,\,n+1} = -\beta(u_n - u_{n+1})$, где $\beta > 0$ — коэффициент квазиупругой силы. Результирующая сила, действующая на n-й атом, $f_n = f_{n,\,n-1} + f_{n,\,n+1} = -\beta(2u_n - u_{n-1} - u_{n+1})$. Уравнение движения (масса $\times$ ускорение $=$ сила) n-го атома имеет вид

$$m\ddot{u}_n + \beta(2u_n - u_{n-1} - u_{n+1}) = 0, \qquad (2.1)$$

где

$$\ddot{u}_n \equiv \frac{d^2u_n}{dt^2}.$$

Выражение для силы f_n может быть получено и другим путем. Потенциальная энергия решетки Φ есть функция от отклонений атомов u_n. Разлагая Φ в ряд по степеням малых отклонений u_n, получим

$$\Phi = \Phi(0) + \sum_n \left(\frac{\partial\Phi}{\partial u_n}\right)_0 u_n + \frac{1}{2}\sum_n\sum_{n'} \left(\frac{\partial^2\Phi}{\partial u_n \partial u_{n'}}\right)_0 u_n u_{n'} + \ldots, \qquad (2.2)$$

где индекс 0 указывает, что все u_n положены равными нулю. Не ограничивая общности, полагаем потенциальную энергию основного состояния $\Phi(0) = 0$. Так как значение $u_n = 0$ соответствует равновесию системы, то $\left(\frac{\partial\Phi}{\partial u_n}\right)_0 = 0$. Для бесконечной атомной цепочки коэффициенты $\left(\frac{\partial^2\Phi}{\partial u_n \partial u_{n'}}\right)_0 = A(|n-n'|)$, т. е. зависят только от расстояния между n-м и n'-м узлами.

По определению сила

$$f_n = -\frac{\partial\Phi}{\partial u_n} = -2\frac{1}{2}\sum_{n'} A(|n-n'|)\,u'_n. \qquad (2.3)$$

Если все $u_{n'} = \text{const} = u_0$, то сила $f_n = 0 = -u_0\sum_{n'} A(|n-n'|)$, где n' пробегает все значения. Если учитывать только ближайших к атому n соседей, то $n' = n,\ n+1,\ n-1$, и мы получим $A(0) + A(1) + A(-1) = 0$. Легко видеть, что (2.3) дает то же выражение для силы f_n, что и (2.1), если положить $A(1) = -\frac{1}{2}A(0) = \beta$.

2. На первый взгляд решение бесконечной системы «зацепляющихся» уравнений (2.1), в которых неизвестная функция $u_n(t)$ связана с «соседними» неизвестными функциями $u_{n-1}(t)$ и $u_{n+1}(t)$, представляет большие трудности. Поэтому удивительно, посредством какого элементарного приема может быть решена

система (2.1). Бесконечная атомная цепочка с квазиупруго взаимодействующими атомами напоминает натянутую струну. Известно, что для бесконечной струны существует простой тип движения в виде бегущей монохроматической волны, для которой отклонение u струны от положения равновесия в точках x в момент t будет: $u(x, t) = A \sin 2\pi \left(\frac{x}{\lambda} - \nu t\right)$, где A — амплитуда, λ — длина волны и ν — частота. Вводя циклическую частоту $\omega = 2\pi\nu$ и волновое число $q = 2\pi/\lambda$, получим $u(x, t) = A \sin(qx - \omega t)$. Если мы условимся рассматривать не только положительные, но и отрицательные значения q, то наряду с волной, распространяющейся вдоль положительной оси x ($q > 0$), получим волны, бегущие в противоположном направлении ($q < 0$). Наконец, если мы учтем, что уравнение колебания струны линейно[1]), так что сумма решений является также интегралом уравнения, то во многих случаях представляется математически более удобным пользоваться комплексной формой решения в виде $u(x, t) = A[\cos(qx - \omega t) + i \sin(qx - \omega t)] = Ae^{i(qx-\omega t)}$, где амплитуда A может быть комплексным числом.

Как мы увидим ниже, бегущая волна в непрерывной струне обладает двумя особенностями, существенно отличающими ее от волн, распространяющихся в дискретных атомных цепочках. Во-первых, абсолютная величина волнового числа q может принимать все значения от 0 до ∞, при этом каждому значению q соответствует определенная форма волны. Во-вторых, частота $\omega = v_0 |q|$ также может изменяться от 0 до ∞.

Попробуем систему (2.1) решить посредством подстановки

$$u_n = Ae^{i(qan - \omega t)}, \tag{2.4}$$

где вместо непрерывной координаты x стоит дискретная величина an (a — расстояние между узлами). Амплитуда же A от номера узла n не зависит. Подставляя (2.4) в (2.1), получим после сокращения обеих частей на $Ae^{i(qan-\omega t)}$

$$-m\omega^2 = -\beta(2 - e^{-iqa} - e^{iqa}).$$

Воспользовавшись соотношением $\cos\alpha = {}^1/_2 (e^{-i\alpha} + e^{+i\alpha})$, получим

$$\left.\begin{aligned} \omega^2 &= 2\frac{\beta}{m}(1 - \cos qa) = 4\frac{\beta}{m}\sin^2\frac{qa}{2} \\ \text{или} \quad \omega &= \omega_m \left|\sin\frac{qa}{2}\right| = \omega_m \left|\sin\frac{\pi a}{\lambda}\right|, \end{aligned}\right\} \tag{2.5}$$

---

[1]) Уравнение свободных колебаний струны имеет вид: $\frac{\partial^2 u}{\partial t^2} = v_0^2 \frac{\partial^2 u}{\partial x^2}$, где $v_0 = \left[\frac{T_0}{\rho}\right]^{1/2}$; T_0 — натяжение, ρ — линейная плотность струны. См. Смирнов В. И. Курс высшей математики — 21 изд. — М.: Наука, 1974, т. 2, § 17, п. 176.

где

$$\omega_m = 2\sqrt{\beta/m}. \qquad (2.5a)$$

Мы видим, что решения (2.4) типа бегущей волны удовлетворяют уравнениям (2.1) для любого n, если частота ω связана с волновым числом q (или длиной волны λ) *соотношением дисперсии* (2.5). Таким образом, в дискретной атомной цепочке имеет место дисперсия, т. е. частота ω не пропорциональна волновому числу q, как это имело место для непрерывной струны.

Заменим в выражении (2.4) q на $q' = q + \frac{2\pi g}{a}$, где g — целое положительное или отрицательное число. Новая волна

$$u_n' = Ae^{i(q'an-\omega t)} = Ae^{i(qan-\omega t)}e^{i2\pi gn} = u_n,$$

так как $\exp(i2\pi gn) = \exp(i2\pi \times \text{целое число}) = 1$.

Таким образом, волна u_n' тождественно (во всех точках и во все моменты времени) совпадает с волной u_n. Но это означает, что q и q' физически неразличимы. Другими словами, достаточно рассматривать изменения q в любом интервале величины $2\pi/a$. При этом все физические свойства нашего одномерного кристалла, зависящие от волнового числа, должны быть периодичны с периодом $2\pi/a$. Мы выберем в качестве основного интервала изменения для q область

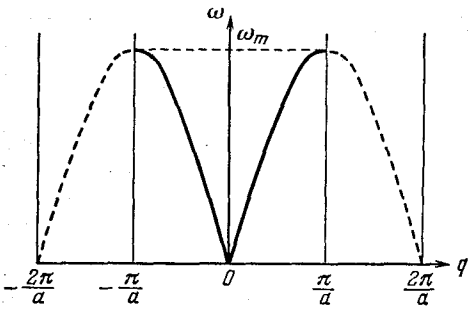


Рис. III.4.

$$-\pi/a \leqslant q \leqslant +\pi/a. \qquad (2.6)$$

На рис. III.4 представлена зависимость ω от q, которая в согласии с вышесказанным периодична с периодом $2\pi/a$. В большинстве случаев достаточно ограничиться рассмотрением положительных значений q от 0 до $q_{\max} = \pi/a$, так как кривая для $q < 0$ симметрична. Максимальному значению q соответствует минимальное значение длины волны λ. Из условия $q_{\max} = 2\pi/\lambda_{\min} = \pi/\lambda$ следует $\lambda_{\min} = 2a$. Представляется вполне наглядным, что в дискретной атомной цепочке не могут существовать волны с $\lambda/2$, меньшей a. Волне с наименьшей $\lambda = \lambda_{\min} = 2a$ соответствует максимальная частота ω_m (2.5а). Существование максимальная частоты также является характерной особенностью колебаний дискретных атомных структур.

3. Макроскопические образцы кристаллов состоят из очень большого, хотя и конечного числа атомов. Если число атомов G

в атомной цепочке очень велико[1]), а силы взаимодействия между атомами распространяются на расстояние одной или нескольких постоянных решетки, то условия, в которых находятся граничные атомы на «поверхности», не влияют на их движение внутри цепочки. В частности, если мы расположим G атомов по окружности очень большого радиуса, так, чтобы G-й и 1-й атомы тоже находились в равновесии на расстоянии a, то граничные условия могут быть заменены условиями цикличности Борна — Кармана, согласно которым

$$u_{n\pm G} = u_n, \tag{2.7}$$

так как в этом случае атом номер $n \pm G$ совпадает с атомом номер n.

Из условий цикличности (2.7) и выражения (2.4) следует, что $(\pm iqaG) = 1$, т. е. $qaG = 2\pi g$, где g — целое число. Отсюда и из (2.6) следует:

$$q = \frac{2\pi}{a}\frac{g}{G}, \tag{2.8}$$

где

$$-G/2 < g < +G/2. \tag{2.8а}$$

Таким образом, для конечной атомной цепочки с G степенями свободы волновое число q, меняющееся в интервале от $-\pi/a$ до $+\pi/a$, принимает G дискретных значений, определяемых неравенством (2.8а). Конечно, мы всегда можем (и даже должны) выбрать G столь большим, чтобы изменение q в (2.8) можно было бы рассматривать как квазинепрерывное. Такая «счетность» q представляется весьма удобной для вопросов статистики и кинетики. При этом конечные формулы, сравнимые с опытом, не должны содержать произвольного числа G атомов основной области.

Так как ω, согласно (2.5), функция q, а q меняется дискретно, то можно поставить вопрос о числе колебаний (с разными q) в интервале частот от ω до $\omega + d\omega$. Из уравнений (2.5) и (2.8) следует, что $d\omega = a\sqrt{\frac{\beta}{m}}\left|\cos\frac{qa}{2}\right| dq$ и $dq = \frac{2\pi}{aG}dg$, откуда $d\omega = \frac{2\pi}{G}\sqrt{\frac{\beta}{m}}\left|\cos\frac{qa}{2}\right| dg$. Число колебаний dz на интервал частоты $d\omega$ в обеих ветвях $\omega(q)$ от $-\pi/a$ до $+\pi/a$ равно

$$dz = 2dg = \frac{G}{\pi}\sqrt{\frac{m}{\beta}}\frac{d\omega}{\left|\cos\frac{aq}{2}\right|}. \tag{2.9}$$

[1]) Как мы увидим ниже, удобно считать G большим нечетным числом, что, очевидно, несущественно.

Из выражения (2.5) следует, что

$$\cos\frac{aq}{2}=\sqrt{1-\sin^2\frac{aq}{2}}=\sqrt{1-\frac{\omega^2}{4\beta/m}};$$

отсюда и из (2.9) получим для плотности числа колебаний на единичный интервал частоты:

$$\frac{dz}{d\omega}=\frac{2G}{\pi}\frac{1}{\sqrt{\omega_m^2-\omega^2}}. \tag{2.10}$$

На рис. III.5 представлена зависимость (2.10). Как мы увидим ниже, для трехмерного кристалла в приближении непрерывного континуума

$$\frac{dz}{d\omega}\sim\omega^2.$$

4. Из теории упругости известно [1]), что скорость распространения звукового импульса в твердом стержне $v_0=\sqrt{E/\rho}$, где

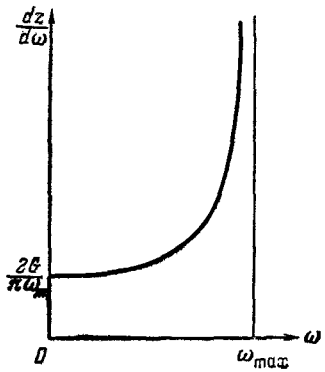


Рис. III.5.

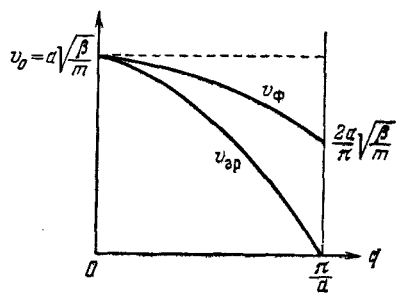


Рис. III.6.

E — модуль Юнга, а ρ — плотность. В случае линейной атомной цепочки $\rho=m/a$,

$$E=\frac{\text{сила}}{\text{относит. удлинение}}=\frac{|f_{n,n-1}|}{\dfrac{|u_n-u_{n-1}|}{a}}=\beta a,$$

откуда

$$v_0=\sqrt{E/\rho}=a\sqrt{\beta/m}. \tag{2.11}$$

Для длинных волн или малых q из (2.5)

$$\omega=2\sqrt{\frac{\beta}{m}}\frac{aq}{2}=v_0q. \tag{2.12}$$

[1]) Хайкин С. Э. Механика./Общий курс физики,—2 изд.—М.: Гостехиздат, 1948, т. 1.

В общем случае, когда имеет место дисперсия волн, т. е. частота ω зависит от волнового числа q, надо различать *фазовую скорость* v_ϕ, с которой распространяется фаза монохроматической волны, и *групповую скорость* $v_{гр}$, с которой распространяется волновой пакет и, следовательно, энергия волн.

Известно [1]), что

$$v_\phi = \omega/|q|, \tag{2.13}$$

$$v_{гр} = |d\omega/dq|. \tag{2.14}$$

Из выражения (2.12) видно, что для длинных волн

$$v_\phi = v_{гр} = v_0 = \text{скорость звука}.$$

В общем случае не малых q, согласно (2.5),

$$v_\phi = v_0 \left| \frac{\sin \frac{aq}{2}}{\frac{aq}{2}} \right|, \tag{2.15}$$

$$v_{гр} = v_0 \left| \cos \frac{aq}{2} \right|. \tag{2.16}$$

На рис. III.6 представлена зависимость v_ϕ и $v_{гр}$ от волнового числа q. Следует отметить, что групповая скорость, с которой переносится энергия колебаний, для наиболее коротких волн падает до нуля.

§ 3. Колебания и волны в сложной одномерной (линейной) решетке

1. В предыдущем параграфе мы рассмотрели движение атомов в одномерной модели простой решетки, когда элементарная ячейка может быть выбрана так, чтобы в ней содержался один атом. Трехмерным аналогом таких решеток являются, например, гранецентрированная и объемноцентрированная кубические решетки, в которых кристаллизуются большинство металлов. Рассмотрим теперь колебания в одномерной модели сложной решетки, когда ее элементарная ячейка содержит два атома. Ионные кристаллы NaCl, CsCl и атомные кристаллы Si и Ge являются примерами решеток, у которых элементарная ячейка содержит 2 атома.

2. На рис. III.7 изображены узлы сложной линейной решетки. Узлы n' заняты атомами с массой m', а узлы n'' — атомами с массой m''. Для общности будем предполагать, что коэффициент квазиупругой силы между атомами n' и n'' равен β_1, а между

[1]) Блохинцев Д. И., § 7.

атомами n' и $n''-1$ равен β_2. Последнее может иметь место и в том случае, когда решетка состоит из атомов одного сорта (в этом случае $m'=m''$). Длина масштабного вектора решетки $\boldsymbol{a}$ равна расстоянию между узлами $n'-1$, n', $n'+1$, ... или узлами $n''-1$, n'', $n''+1$, ... Поэтому элементарная ячейка «объема» $\Omega_0=a$ содержит два атома. Обозначая смещения n'-го и n''-го атомов через u'_n и u''_n, получим при учете взаимодействия

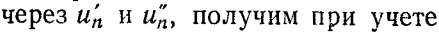


Рис. III.7.

только ближайших атомов, в приближении квазиупругой силы, следующие уравнения движения (см. § 2, п. 1):

$$\begin{aligned} m'\ddot{u}'_n &= -\beta_1(u'_n-u''_n)-\beta_2(u'_n-u''_{n-1}),\\ m''\ddot{u}''_n &= -\beta_1(u''_n-u'_n)-\beta_2(u''_n-u'_{n+1}). \end{aligned} \tag{3.1}$$

Попытаемся удовлетворить этой системе уравнений, полагая

$$u'_n=A'e^{i(qan-\omega t)},\quad u''_n=A''e^{i(qan-\omega t)}, \tag{3.2}$$

где q и ω — волновое число и частота (одинаковые в обоих выражениях): а $A'\neq A''$ — амплитуды.

Подставляя (3.2) в уравнения (3.1), получим после простых преобразований и сокращения обеих частей равенств на множитель $e^{i(qan-\omega t)}$:

$$\left[\omega^2-\frac{\beta_1+\beta_2}{m'}\right]A'+\left[\frac{\beta_1+\beta_2e^{-iaq}}{m'}\right]A''=0, \tag{3.3}$$

$$\left[\frac{\beta_1+\beta_2e^{iaq}}{m''}\right]A'+\left[\omega^2-\frac{\beta_1+\beta_2}{m''}\right]A''=0. \tag{3.3а}$$

Мы получили систему линейных однородных уравнений относительно неизвестных, вообще говоря, комплексных амплитуд A' и A''. Очевидно, что из такой системы можно определить только отношение A'/A'', так как cA' и cA'' ($c=\text{const}$) также удовлетворяют уравнениям (3.3) и (3.3а). Для совместного решения (3.3) и (3.3а) необходимо, чтобы отношение A'/A'', определенное из каждого уравнения, было одинаково, т. е.

$$\frac{A'}{A''}=\frac{\beta_1+\beta_2e^{-iaq}}{(\beta_1+\beta_2)-m'\omega^2}=\frac{(\beta_1+\beta_2)-m''\omega^2}{\beta_1+\beta_2e^{iaq}}. \tag{3.4}$$

Полученная пропорция (3.4) приводит к алгебраическому

квадратному уравнению для ω^2. Корни этого квадратного уравнения [1])

$$\omega_1^2 = \frac{1}{2}\omega_0^2\left\{1 - \sqrt{1-\gamma^2\sin^2\frac{aq}{2}}\right\},$$
$$\omega_2^2 = \frac{1}{2}\omega_0^2\left\{1 + \sqrt{1-\gamma^2\sin^2\frac{aq}{2}}\right\}, \tag{3.5}$$

где $\omega_0^2 = \frac{(\beta_1+\beta_2)(m'+m'')}{m'm''}$ и $\gamma^2 = 16\left[\frac{\beta_1\beta_2}{(\beta_1+\beta_2)^2}\right]\left[\frac{m'm''}{(m'+m'')^2}\right]$. Величина γ^2 достигает своего наибольшего значения, равного единице, при $\beta_1=\beta_2$ и $m'=m''$. Отсюда видно, что $\gamma^2\sin^2\frac{aq}{2}\leqslant 1$, что обеспечивает вещественность частот ω_1 и ω_2.

Мы видим, что аналогично случаю простой линейной цепочки (§ 2) решение (3.2) удовлетворяет уравнениям движения (3.1), если ω и q связаны законом дисперсии (3.5). Однако имеется и важное различие по сравнению с предыдущим случаем. Так, выражения (3.5) определяют две ветви дисперсии, одну из которых $\omega_1=\omega_{ак}$ мы по причинам, указанным ниже, будем называть *акустической*, другую $\omega_2=\omega_{оп}$ — *оптической*.

Рассуждая так же, как в п. 2 предыдущего параграфа, можно показать, что характер решения (3.2) позволяет ограничиться при рассмотрении изменения q интервалом $(0, \pi/a)$. Из выражения (3.5) следует, что для $q=0$, π/a:

$$\omega_{ак}(0)=0, \quad \omega_{ак}(\pi/a)=\frac{\omega_0}{\sqrt{2}}\sqrt{1-\sqrt{1-\gamma^2}},$$
$$\omega_{оп}(0)=\omega_0, \quad \omega_{оп}(\pi/a)=\frac{\omega_0}{\sqrt{2}}\sqrt{1+\sqrt{1-\gamma^2}}. \tag{3.6}$$

Отсюда видно, что

$$\omega_{оп}(0)=\omega_0>\omega_{оп}(\pi/a)>\omega_{ак}(\pi/a)>\omega_{ак}(0)=0. \tag{3.6а}$$

Для длинных волн $\lambda\gg a$ или $aq\ll 1$ из выражений (3.5) в приближении $\sin\frac{aq}{2}\approx\frac{aq}{2}$, следует, при разложении корней в ряд:

$$\omega_{ак}\approx\frac{1}{4}\omega_0\gamma aq\sim q, \quad \omega_{оп}\approx\omega_0\left(1-\frac{\gamma^2a^2}{32}q^2\right). \tag{3.7}$$

Беря производную по q от $\omega_{1,2}$ (см. (3.5)), легко показать что

$$\left(\frac{d\omega_{ак}}{dq}\right)_{\pi/a}=\left(\frac{d\omega_{оп}}{dq}\right)_{\pi/a}=0,$$

если $\gamma^2\neq 1$. Используя это обстоятельство и равенства (3.6) и (3.7), мы приходим к однозначному ходу дисперсионных ветвей,

[1]) При алгебраических преобразованиях мы пользовались соотношениями: $e^{iaq}+e^{-iaq}=2\cos aq$ и $1-\cos aq=2\sin^2\frac{aq}{2}$.

изображенному на рис. III.8, *а*. Такая картина имеет место, когда $\gamma^2 < 1$, а для этого достаточно, чтобы или $m' \neq m''$, или $\beta_1 \neq \beta_2$. Подчеркиваем еще раз существенную особенность акустической и оптической ветвей: при $\lambda \to \infty$ частота $\omega_{ак} \to 0$, а $\omega_{оп} \to \omega_0 \neq 0$.

Рассмотрим «вырожденный случай» (рис. III.8, *б*), когда $m' = m'' = m$ и $\beta_1 = \beta_2 = \beta$, т. е. $\gamma^2 = 1$; из выражений (3.5) следует:

$$\omega_{ак} = 2\sqrt{\frac{\beta}{m}}\left|\sin\frac{aq}{4}\right|, \quad \omega_{оп} = 2\sqrt{\frac{\beta}{m}}\left|\cos\frac{aq}{4}\right|. \qquad (3.8)$$

В этом случае тождественности атомов и одинаковости связей расстояния между всеми атомами должны быть равны, и, следовательно, постоянная решетки в наших обозначениях равна

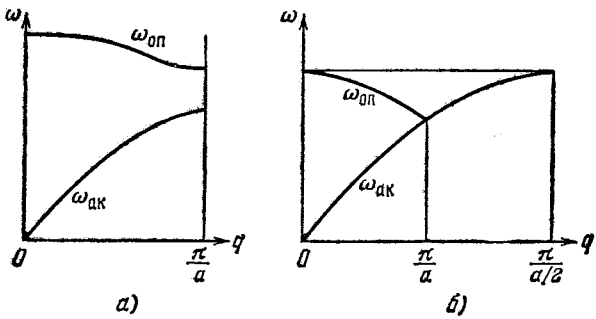


Рис. III.8.

$a/2$, т. е. вдвое меньше старой. Мы должны теперь рассматривать изменение волнового числа q в интервале $\left(0, \frac{\pi}{a/2}\right)$. Легко видеть, что $\omega_{ак}$ в (3.8) соответствует $\omega(q)$ в выражении (2.5) для интервала $\left(0, \frac{1}{2}\frac{\pi}{a/2}\right)$, а $\omega_{оп}$ — также $\omega(q)$ в выражении (2.5) для интервала $\left(\frac{1}{2}\frac{\pi}{a/2}, \frac{\pi}{a/2}\right)$ с эквивалентным значением q, равным $q' = \frac{2\pi}{a} - q$; в самом деле,

$$\cos\frac{aq'}{4} = \cos\left(\frac{\pi}{2} - \frac{aq}{4}\right) = \sin\frac{aq}{4}.$$

3. Исследуем характер колебаний атомов в акустической и оптической ветвях. Из соотношений (3.2) и (3.4) следует:

$$\frac{u_n'}{u_n''} = \frac{A'}{A''} = \frac{\beta_1 + \beta_2 e^{-iqa}}{(\beta_1 + \beta_2) - m'\omega^2}. \qquad (3.9)$$

Отметим, что для $q = 0$ и $q = \pi/a$ множитель $e^{-iqa} = 1$ и $e^{-iqa} = -1$. В обоих случаях, как это видно из равенства (3.9), амплитуды A' и A'' и смещения u_n' и u_n'' можно считать вещественными.

Рассмотрим вначале важный случай длинных (строго говоря, бесконечно длинных) волн:

$$\lambda = \infty, \quad q = \frac{2\pi}{\lambda} = 0.$$

В этом случае $e^{-iqa} = e^0 = 1$. Используя выражения (3.6), т. е. предельные значения частоты ω для $q = 0$, получим из (3.9)

$$\left(\frac{u_n'}{u_n''}\right)_{\text{ак}} = 1, \quad \left(\frac{u_n'}{u_n''}\right)_{\text{оп}} = -\frac{m''}{m'}. \tag{3.10}$$

Таким образом, в бесконечно длинной волне акустической ветви колебаний атомы движутся синхронно и отклонения в каждый момент времени одинаковы ($u_n' = u_n''$). В бесконечно длинной оптической ветви колебаний атомы ячейки колеблются в противоположных фазах, так что их центр тяжести остается неподвижным ($m'u_n' + m''u_n'' = 0$). Движение атомов в первом случае соответствует тому, что имеет место при распространении упругих акустических волн, чем и объясняется название соответствующей ветви колебаний. Если ячейка сложного кристалла состоит из разноименных ионов, то колебания второго рода связаны с изменением электрического дипольного момента ячейки и часто являются оптически активными, т. е. проявляются в поглощении и испускании инфракрасного излучения. Этим объясняется, что вторая ветвь колебаний называется оптической.

Рассмотрим теперь случай коротких волн:

$$\lambda = 2a, \quad q = 2\pi/\lambda = \pi/a.$$

В этом случае $e^{-iaq} = e^{-i\pi} = -1$. Используя выражения для $\omega(\pi/a)$ (см. (3.6)), получим из равенства (3.9)

$$\frac{u_n'}{u_n''} = \frac{\dfrac{(\beta_1 - \beta_2)}{(\beta_1 + \beta_2)}}{1 - \dfrac{m' + m''}{2m''}\left[1 \mp \sqrt{1 - \dfrac{\beta_1\beta_2}{(\beta_1+\beta_2)^2}\,\dfrac{m'm''}{(m'+m'')^2}}\right]}, \tag{3.11}$$

где верхний знак перед корнем соответствует акустической, а нижний знак — оптической ветви колебаний. В данном случае представляется целесообразным рассмотреть ряд частных случаев.

Если $\beta_1 = \beta_2$, то числитель в (3.11) равен нулю, а знаменатель отличен от нуля и равен $(m'' - m')/m''$ для акустической ветви (в случае $m' < m''$), а для оптической — в случае $m' > m''$.

Отсюда можно заключить, что

$$\begin{aligned} &u_n' = 0, \ u_n'' \neq 0 \text{ для акустической ветви в случае } m' < m'', \\ &u_n' = 0, \ u_n'' \neq 0 \text{ для оптической ветви в случае } m' > m''. \end{aligned} \tag{3.12}$$

Для предельно коротких волн $\lambda = 2a$ в акустической ветви неподвижны более легкие атомы m', а колеблются более тяжелые m''. В оптической ветви колебаний — наоборот.

Мы рассмотрели для акустической ветви случай $m' < m''$ для оптической — случай $m' > m''$. Сделали мы это только для удобства. Если бы мы предположили для акустической ветви (верхний знак перед корнем в (3.11), что $m' > m''$, то знаменатель оказался бы равным нулю и $\left(\frac{u_n'}{u_n''}\right) = \frac{0}{0}$. Раскрыв неопределенность, мы убедились бы, что в этом случае $\left(\frac{u_n'}{u_n''}\right) = \infty$, т. е. $u_n' \neq 0$ и $u_n'' = 0$. Но это привело бы нас к прежнему физическому заключению, что в акустической ветви неподвижны более легкие атомы m'', а колеблются более тяжелые m'.

Рассмотрим теперь второй частный случай: $m' = m''$ и $\beta_1 > \beta_2$. Знаменатель в (3.11) равен $1 - 1 \pm \frac{\sqrt{(\beta_1-\beta_2)^2}}{\beta_1+\beta_2} = \pm \frac{\sqrt{(\beta_1-\beta_2)^2}}{\beta_1+\beta_2}$, т.е. отношению $(\beta_1-\beta_2)/(\beta_1+\beta_2)$ — для акустической ветви и для оптической ветви $-(\beta_1-\beta_2)/(\beta_1+\beta_2)$. Таким образом,

$$\left(\frac{u_n'}{u_n''}\right)_{\text{ак}} = 1 \quad \text{и} \quad \left(\frac{u_n'}{u_n''}\right)_{\text{оп}} = -1, \tag{3.13}$$

т. е. в этом случае атомы в предельно короткой акустической волне колеблются в фазе, а в предельно короткой оптической волне в противофазе. Аналогично можно рассмотреть случай $\beta_1 < \beta_2$; тогда правые части равенств (3.13) меняются местами.

4. Аналогично тому, как было сделано в § 2 п. 3, можно и в случае сложной решетки ввести циклические условия Борна—Кармана (2.7), где теперь G — число ячеек основной области. Можно определить плотность числа колебаний $\frac{dz}{d\omega}$ для акустической и оптической ветвей подобно тому, как мы сделали в (2.10).

§ 4. Нормальные координаты для простой одномерной решетки

В § 2 мы рассмотрели колебания и волны в линейной цепочке из одинаковых атомов и показали, что выражение (2.4) удовлетворяет уравнениям движения (2.1), если выполнены соотношения дисперсии (2.5). Очевидно, что гармонические волны (2.4) не описывают наиболее общего движения атомов цепочки. Произвольное движение атомов может быть представлено посредством линейной суперпозиции (суммы) всевозможных волн типа (2.4). Отдельные волны будут при этом отличаться волновым числом q, соответствующей ему частотой $\omega(q) \equiv \omega_q$ и амплитудой A_q, обычно различной у разных волн. Положим реальные вещественные смещения атомов цепочки

$$u_n = A_q e^{i(qan-\omega_q t)} + A_q^* e^{-i(qan-\omega_q t)} = 2|A_q|\cos(qan - \omega_q t + \alpha_q),$$

где $A_q = |A_q| e^{i\alpha} q$.

Таким образом, в наиболее общем случае движения атомов

$$u_n = \sum_q \{A_q e^{i(qan-\omega)} + A_q^* e^{-i(qan-\omega_q t)}\} = \frac{1}{\sqrt{G}} \sum_q \{a_q e^{iqan} + a_q^* e^{-iqan}\}, \tag{4.1}$$

где $a_q \equiv \sqrt{G} A_q e^{-i\omega_q t}$. При этом мы предполагаем, что на решетку наложено условие цикличности, так что суммирование в (4.1) производится по G дискретным значениям $q = \frac{2\pi}{aG} g$ (см. (2.8)).

Как мы увидим ниже, величины a_q или, точнее, некоторые простые комбинации из них, являются нормальными координатами и соответствующими обобщенными импульсами нашей линейной атомной цепочки (Приложение 5). Таким образом, соотношение (4.1) можно рассматривать как преобразование от координат u_n к нормальным координатам и импульсам a_q (величины a_q комплексны, поэтому им соответствует $2G$ вещественных величин). Это преобразование имеет несколько более общий характер, чем рассмотренное в Приложении 5, так как в (4.1) старые координаты u_n выражаются одновременно через новые нормальные координаты и соответствующие им импульсы (касательное или контактное преобразование[1])). Кинетическая энергия $\mathscr{T}$ и потенциальная Φ для атомной цепочки равны

$$\mathscr{T} = \frac{m}{2} \sum_{n=1}^{G} \dot{u}_n^2, \tag{4.2}$$

$$\Phi = \frac{\beta}{2} \sum_{n=1}^{G} (u_n - u_{n-1})^2. \tag{4.3}$$

В самом деле, из (4.3) следует, что сила, действующая на n-ю частицу,

$$f_n = -\frac{\partial \Phi}{\partial u_n} = -\beta (2u_n - u_{n-1} - u_{n+1}),$$

совпадает с (2.1).

Используя (4.1) и учитывая, что $\dot{a}_q = -i\omega_q a_q$ и $\dot{a}_q^* = i\omega_q a_q^*$, получим, что кинетическая энергия

$$\mathscr{T} = \frac{m}{2} \sum_{n=1}^{G} \dot{u}_n \dot{u}_n = \frac{m}{2} \sum_{n=1}^{G} \frac{1}{\sqrt{G}} \sum_q \{\dot{a}_q e^{iqan} + \dot{a}_q^* e^{-iqan}\} \times$$

$$\times \frac{1}{\sqrt{G}} \sum_{q'} \{\dot{a}_{q'} e^{iq'an} + \dot{a}_{q'}^* e^{-iq'an}\} = -\frac{m}{2G} \sum_{qq'} \sum_{n=1}^{G} \omega_q \omega_{q'} \{a_q a_{q'} e^{i(q+q')an} -$$

$$- a_q a_{q'}^* e^{i(q-q')an} - a_q^* a_{q'} e^{-i(q-q')an} + a_q^* a_{q'}^* e^{-i(q+q')an}\}. \tag{4.4}$$

[1]) Ландау Л. Д., Лифшиц Е. М./Теоретическая физика, т. 1. Механика.— 4 изд.— М.: Наука, 1973, гл. VII.

Легко показать (см. Приложение 6), что

$$\sum_{n=1}^{G} e^{iqan} = \sum_{n=1}^{G} e^{i\frac{2\pi}{G}gn} = \begin{cases} 0, & \text{когда } q \neq 0 \text{ (или } g \neq 0), \\ G, & \text{когда } q = 0 \left(\text{или } \frac{2\pi}{a} \times \text{целое число}\right). \end{cases} \tag{4.5}$$

Таким образом, при суммировании по n первого слагаемого в фигурной скобе (4.4) получим

$$\sum_{n=1}^{G} e^{i(q+q')an} = \begin{cases} 0 & \text{при } q+q' \neq 0, \\ G & \text{при } q+q' = 0 \text{ (или } q' = -q). \end{cases} \tag{4.6}$$

Аналогично суммируются по n второе, третье и четвертое слагаемое в фигурной скобке в (4.4). Учитывая, что, согласно (2.5), $\omega_{-q} = \omega_q$, получим

$$\mathcal{T} = \frac{m}{2} \sum_q \omega_q^2 (2a_q a_q^* - a_q a_{-q} - a_q^* a_{-q}^*). \tag{4.7}$$

Выразим теперь потенциальную энергию (4.3) через величины a_q и a_q^*. Подставляя в (4.3) u_n и u_{n-1} из (4.1), получим

$$\Phi = \frac{\beta}{2} \sum_{n=1}^{G} (u_n - u_{n-1})(u_n - u_{n-1}) = \frac{\beta}{2G} \sum_{qq'} \sum_{n=1}^{G} [a_q e^{iqan} + a_q^* e^{-iqan} - a_q e^{-iqa} e^{iqan} - a_q^* e^{iqa} e^{-iqan}][a_{q'} e^{iq'an} + a_{q'}^* e^{-iq'an} - a_{q'} e^{-iq'a} e^{iq'an} - a_{q'}^* e^{iq'a} e^{-iq'an}].$$

Перемножая квадратные скобки и группируя слагаемые, получим

$$\sum_{n=1}^{G} \{a_q a_{q'} (1 + e^{-i(q+q')a} - e^{-iqa} - e^{-iq'a}) e^{i(q+q')an} + a_q a_{q'}^* (1 + e^{i(q'-q)a} - e^{iq'a} - e^{-iqa}) e^{i(q-q')an} + a_q^* a_{q'} (1 + e^{i(q-q')a} - e^{-iq'a} - e^{iqa}) e^{i(q'-q)an} + a_q^* a_{q'}^* (1 + e^{i(q+q')a} - e^{iqa} - e^{iq'a}) e^{-i(q+q')an}\}.$$

Учитывая условия (4.6), видим, что при суммировании по n мы получим результат, отличный от нуля и равный G только в тех случаях, когда $q' = -q$ или $q' = +q$. При этом все круглые скобки становятся равными

$$2 - e^{-iqa} - e^{iqa} = 2(1 - \cos qa) = 4 \sin^2 \frac{aq}{2} = \frac{m\omega_q^2}{\beta},$$

как это следует из (2.5). Таким образом,

$$\Phi = \frac{m}{2} \sum_q \omega_q^2 (2a_q a_q^* + a_q a_{-q} + a_q^* a_{-q}^*). \tag{4.8}$$

Полная энергия

$$\mathcal{E} = \mathcal{T} + \Phi = 2m \sum \omega_q^2 a_q a_q^*. \tag{4.9}$$

Определим величины x_q и p_q посредством равенств

$$x_q = a_q + a_q^* = 2\mathrm{Re}\{a_q\}, \quad p_q = \frac{m\omega_q}{i}(a_q - a_q^*) = \frac{m\omega_q}{i} 2\mathrm{Im}\{a_q\}. \quad (4.10)$$

Отсюда

$$a_q = \frac{1}{2}\left(x_q + i\frac{p_q}{m\omega_q}\right), \quad a_q^* = \frac{1}{2}\left(x_q - i\frac{p_q}{m\omega_q}\right). \quad (4.10a)$$

Подставляя эти значения в выражение (4.9), получим

$$\mathscr{E} = \sum_q \left\{\frac{1}{2m}p_q^2 + \frac{1}{2}m\omega_q^2 x_q^2\right\} = \mathscr{H}(x_q, p_q). \quad (4.11)$$

Мы видим, что величины x_q и $p_q = m\dot{x}_q$ играют роль нормальных координат и сопряженных им импульсов. Выражение (4.11) представляет собой функцию Гамильтона, выраженную в этих переменных. Таким образом, полная энергия $\mathscr{E}$ наиболее общего движения атомов одномерного кристалла может быть представлена как сумма энергий нормальных колебаний ведущих себя подобно линейным гармоническим осцилляторам с частотами ω_q.

§ 5. Колебания атомов трехмерной сложной кристаллической решетки

1. В этом параграфе мы покажем, что многие особенности колебаний одномерной (линейной) атомной цепочки имеют место и в случае трехмерной кристаллической решетки.

Рассмотрим сложный кристалл с s различными (в общем случае) атомами в элементарной ячейке, имеющими массы m_k ($k = 1, 2, \ldots, s$). В дальнейшем мы будем рассматривать основную область кристалла объема V, имеющую форму параллелепипеда, построенного на векторах $G\boldsymbol{a}_i$ ($i = 1, 2, 3$), где $\boldsymbol{a}_i$ — трансляционные векторы прямой решетки, а G — большое нечетное число. Очевидно, что $V = G^3(\boldsymbol{a}_1[\boldsymbol{a}_2\boldsymbol{a}_3]) = G^3\Omega_0 = N\Omega_0$, где Ω_0 — объем элементарной ячейки и $N = G^3$. Положение атома k-го сорта в $\boldsymbol{n}$-й элементарной ячейке определяется радиусом-вектором

$$\boldsymbol{r}_n^k = \boldsymbol{a}_n + \boldsymbol{r}^k, \quad (5.1)$$

где $\boldsymbol{a}_n = \boldsymbol{n}_1\boldsymbol{a}_1 + \boldsymbol{n}_2\boldsymbol{a}_2 + \boldsymbol{n}_3\boldsymbol{a}_3$ — вектор прямой решетки, а $\boldsymbol{r}^k$ — радиус-вектор, определяющий положение k-го атома внутри кристаллической ячейки.

Обозначим через $\boldsymbol{u}_n^k$ смещение k-го атома $\boldsymbol{n}$-й ячейки из положения равновесия, а прямоугольные проекции этого смещения — через $u_{n\alpha}^k$ ($\alpha = x, y, z$).

Потенциальная энергия Φ основной области кристалла является функцией $3sN$ смещений $u_{n\alpha}^k$ и, очевидно, обладает минимумом при $u_{n\alpha}^k = 0$; так что $\left(\frac{\partial\Phi}{\partial u_{n\alpha}^k}\right)_0 = 0$.

Разложим Φ в ряд по степеням проекций смещений $u^k_{n\alpha}$:

$$\Phi=\frac{1}{2}\sum_{nn'kk'\alpha\beta}\Phi_{\alpha\beta}\binom{kk'}{nn'}u^k_{n\alpha}u^{k'}_{n'\beta}+\frac{1}{6}\sum_{\substack{nn'n''kk'k''\\ \alpha\beta\gamma}}\Phi_{\alpha\beta\gamma}\binom{kk'k''}{nn'n''}u^k_{n\alpha}u^{k'}_{n'\beta}u^{k''}_{n''\gamma}. \tag{5.2}$$

Мы положили потенциальную энергию в минимуме равной нулю и ввели следующие обозначения:

$$\Phi_{\alpha\beta}\binom{kk'}{nn'}=\left(\frac{\partial^2\Phi}{\partial u^k_{n\alpha}\partial u^{k'}_{n'\beta}}\right)_{00},\quad \Phi_{\alpha\beta\gamma}\binom{kk'k''}{nn'n''}=\left(\frac{\partial^3\Phi}{\partial u^k_{n\alpha}\partial u^{k'}_{n'\beta}\partial u^{k''}_{n''\gamma}}\right)_{000}. \tag{5.2а}$$

Можно показать, что коэффициенты разложения (5.2а) удовлетворяют соотношениям:

$$\Phi_{\alpha\beta}\binom{kk'}{nn'}=\Phi_{\alpha\beta}\binom{kk'}{n-n'};\quad \Phi_{\alpha\beta\gamma}\binom{kk'k''}{nn'n''}=\Phi_{\alpha\beta\gamma}\binom{kk'k''}{n-n',\ n-n''}, \tag{5.3}$$

$$\Phi_{\alpha\beta}\binom{kk'}{nn'}=\Phi_{\beta\alpha}\binom{k'k}{n'n}, \tag{5.3а}$$

$$\sum_{n'k'}\Phi_{\alpha\beta}\binom{kk'}{nn'}=0,\quad \sum_{n'n''k'k''}\Phi_{\alpha\beta\gamma}\binom{kk'k''}{nn'n''}=0. \tag{5.3б}$$

Равенства (5.3) следуют из того, что коэффициенты, определяющие взаимодействие атомов, зависят только от их сорта (k, k') и от расстояния между соответствующими ячейками ($n-n'$, $n-n''$). (5.3а) непосредственно вытекает из определения (5.2а). Для доказательства (5.3б) разложим в ряд по смещениям атомов функцию $\partial\Phi/\partial u^k_{n\alpha}$:

$$\frac{\partial\Phi}{\partial u^k_{n\alpha}}=\left(\frac{\partial\Phi}{\partial u^k_{n\alpha}}\right)_0+\sum_{n'k'\beta}\left(\frac{\partial^2\Phi}{\partial u^k_{n\alpha}\partial u^{k'}_{n'\beta}}\right)_{00}u^{k'}_{n'\beta}+$$
$$+\frac{1}{2}\sum_{\substack{n'k'\beta\\ n''k''\gamma}}\left(\frac{\partial^3\Phi}{\partial u^k_{n\alpha}\partial u^{k'}_{n'\beta}\partial u^{k''}_{n''\gamma}}\right)_{000}u^{k'}_{n'\beta}u^{k''}_{n''\gamma}. \tag{5.4}$$

Пусть все атомы испытали одинаковое смещение, равное u^0; тогда кристалл сместился в пространстве как целое и изменение всех функций, зависящих от взаимного расположения атомов, равно нулю, т. е. $\partial\Phi/\partial u^k_{n\alpha}=0$ (из условий равновесия $\left(\frac{\partial\Phi}{\partial u^k_{n\alpha}}\right)_0=0$); таким образом, из (5.4) следует:

$$0=\sum_\beta u^0_\beta\sum_{n'k'}\Phi_{\alpha\beta}\binom{kk'}{nn'}+\frac{1}{2}\sum_{\beta\gamma}u^0_\beta u^0_\gamma\sum_{n'n''k'k''}\Phi_{\alpha\beta\gamma}\binom{kk'k''}{nn'n''}=0.$$

Отсюда, ввиду независимости проекций u^0_β и u^0_γ, следуют

соотношения (5.3б). Кинетическая энергия движения атомов равна

$$\mathscr{T} = \frac{1}{2}\sum_{nk} m_k (\dot{u}_n^k)^2 = \frac{1}{2}\sum_{nk\alpha} m_k (\dot{u}_{n\alpha}^k)^2. \tag{5.5}$$

2. Классические уравнения движения атомов в гармоническом приближении имеют вид

$$m_k \ddot{u}_{n\alpha}^k = -\frac{\partial \Phi}{\partial u_{n\alpha}^k} = -\sum_{n'k'\beta} \Phi_{\alpha\beta}\begin{pmatrix} kk' \\ nn' \end{pmatrix} u_{n'\beta}^{k'} \tag{5.6}$$

$$(n = 1, 2, 3, \ldots, N;\ k = 1, 2, 3, \ldots, s;\ \alpha = x, y, z),$$

т. е. совокупной системы $3sN$ дифференциальных уравнений для $3sN$ неизвестных функций $u_{n\alpha}^k(t)$.

Аналогично сложной линейной решетке (см. (3.2)) ищем решение системы (5.6) в виде бегущих волн:

$$\tilde{u}_{n\alpha}^k = \frac{1}{\sqrt{m_k}} A_\alpha^k(q)\, e^{i(qa_n - \omega t)}, \tag{5.7}$$

где $\frac{1}{\sqrt{m_k}} A_\alpha^k (\alpha = x, y, z)$ проекция комплексной амплитуды $\frac{1}{\sqrt{m_k}}\boldsymbol{A}^k$, различной для разных сортов атомов ($k = 1, 2, \ldots, s$), $\boldsymbol{q} = \frac{2\pi}{\lambda}\boldsymbol{\nu}$ — волновой вектор ($\boldsymbol{\nu}$ — единичный вектор нормали к плоской волне, длина которой равна λ) и $\omega = \omega(\boldsymbol{q}) \equiv \omega_q$ — циклическая частота. Величины $\tilde{u}_{n\alpha}^k$ комплексны, так что реальные смещения $u_{n\alpha}^k = \mathrm{Re}\,(\tilde{u}_{n\alpha}^k)$. Мы можем искать решение уравнений (5.6) в комплексной форме (5.7), состоящей из суммы косинуса и синуса в силу линейности (5.6). Так же как и в одномерном случае, волновой вектор $\boldsymbol{q}$ обладает свойствами, обусловленными тем, что плоская волна (5.7) распространяется в дискретной системе (решетке). Заменим в (5.7) $\boldsymbol{q}$ на $\boldsymbol{q}' = \boldsymbol{q} + \boldsymbol{b}_g$, где $\boldsymbol{b}_g = g_1\boldsymbol{b}_1 + g_2\boldsymbol{b}_2 + g_3\boldsymbol{b}_3$ — вектор обратной решетки; тогда $\boldsymbol{q}'\boldsymbol{a}_n = \boldsymbol{q}\boldsymbol{a}_n + 2\pi(n_1g_1 + n_2g_2 + n_3g_3) = \boldsymbol{q}\boldsymbol{a}_n + 2\pi k$ (k — целое число). Таким образом, из (5.7) следует:

$$(\tilde{u}_{n\alpha}^k)' = \frac{1}{\sqrt{m_k}} A_\alpha^k e^{i(q'a_n - \omega t)} = \frac{1}{\sqrt{m_k}} A_\alpha^k e^{i(qa_n - \omega t)} e^{i2\pi k} = \tilde{u}_{n\alpha}^k, \tag{5.8}$$

так как $e^{i2\pi k} = 1$. Из (5.8) видно, что волна с волновым вектором $\boldsymbol{q}'$ совпадает с волной с вектором $\boldsymbol{q}$. Но это означает, что вектор $\boldsymbol{q}'$ физически эквивалентен вектору $\boldsymbol{q}$. Это позволяет, так же как и в одномерном случае, рассматривать изменения волнового вектора $\boldsymbol{q}$ в ограниченной области. Полагая $\boldsymbol{a}_n = \boldsymbol{a}_i$ и $\boldsymbol{b}_g = \boldsymbol{b}_i$, получим $\boldsymbol{q}'\boldsymbol{a}_i = (\boldsymbol{q} + \boldsymbol{b}_i)\boldsymbol{a}_i = \boldsymbol{q}\boldsymbol{a}_i + 2\pi$. Отсюда видно,

что величина $\boldsymbol{qa}_i$ всегда может быть выбрана в интервале 2π. Мы положим

$$-\pi \leqslant \boldsymbol{qa}_i \leqslant +\pi \quad (i=1, 2, 3)^{1}). \tag{5.9}$$

Для кубического кристалла

$$-\frac{\pi}{a} \leqslant q_\alpha \leqslant +\frac{\pi}{a} \quad (\alpha = x, y, z), \tag{5.9а}$$

если прямоугольные оси координатной системы направить по ребрам кубической ячейки.

Для произвольной решетки область изменения $\boldsymbol{qa}_i$, равная 2π, может быть выбрана в виде элементарной ячейки обратной решетки, т. е. в форме параллелепипеда с ребрами $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ и $\boldsymbol{b}_3$, или в симметричной форме в виде первой бриллюэновской зоны (гл. II, § 9, п. 1). Подставляя (5.7) в систему (5.6) и деля обе части уравнения на множитель $e^{i(\boldsymbol{qa}_n - \omega t)}$, получим

$$\omega^2 A_\alpha^k = \sum_{k'\beta} D_{\alpha\beta}^{kk'} A_\beta^{k'}, \tag{5.10}$$

где элементы динамической матрицы кристалла

$$D_{\alpha\beta}^{kk'}(\boldsymbol{q}) = \sum_{n'} \frac{1}{\sqrt{m_k m_{k'}}} \Phi_{\alpha\beta}\begin{pmatrix} kk' \\ nn' \end{pmatrix} e^{iq\left(a'_n - a_n\right)}. \tag{5.10а}$$

В этой матрице $3s$-го ранга двойной индекс $\begin{pmatrix} k \\ \alpha \end{pmatrix}$ нумерует строки, а индекс $\begin{pmatrix} k' \\ \beta \end{pmatrix}$ столбцы. Мы получили линейную однородную алгебраическую систему $3s$ уравнений (вместо $3sN$ уравнений (5.6)) для $3s$ неизвестных комплексных величин A_α^k ($k=1, 2, \ldots, s$; $\alpha = x, y, z$). Система (5.10) может быть переписана в более компактной форме:

$$\sum_{k'\beta} \{D_{\alpha\beta}^{kk'} - \omega^2 \delta_{kk'}\delta_{\alpha\beta}\} A_\beta^{k'} = 0 \quad (k=1, 2, \ldots, s;\ \alpha = x, y, z), \tag{5.10б}$$

где символы $\delta_{kk'}$ и $\delta_{\alpha\beta}$, отличные от нуля и равные единице только при $k=k'$ и $\alpha=\beta$, автоматически обеспечивают совпадение системы (5.10б) с (5.10).

Известно[2]), что система (5.10б) имеет решения, отличные от нулевых (тривиальных) только в том случае, если определитель системы равен нулю:

$$|D_{\alpha\beta}^{kk'} - \omega^2 \delta_{kk'}\delta_{\alpha\beta}| = 0. \tag{5.11}$$

---

[1]) Полезно напомнить, что более глубокой причиной появления волнового вектора $\boldsymbol{q}$ и его свойств является трансляционная симметрия кристаллической решетки (см. гл. II, § 9).

[2]) Смирнов В. И., Курс высшей математики.— 10 изд.— М.: Наука, 1974, т. 3, ч. I., п. 10.

Это *характеристическое уравнение* степени $3s$ для частоты ω^2 может быть расписано подробнее:

$$\begin{vmatrix} D_{xx}^{11}-\omega^2, & D_{xy}^{11}, & D_{xz}^{11}, & D_{xx}^{12}, \ldots, & D_{xz}^{1s} \\ D_{yx}^{11}, & D_{yy}^{11}-\omega^2, & D_{yz}^{11}, & D_{yx}^{12}, \ldots, & D_{yz}^{1s} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ D_{zx}^{s1}, & D_{zy}^{s1}, & D_{zz}^{s1}, & D_{zx}^{s2}, \ldots, & D_{zz}^{ss}-\omega^2 \end{vmatrix} = 0. \qquad (5.11а)$$

Из (5.10а) и (5.3а) следует, что

$$D_{\alpha\beta}^{kk'} = D_{\beta\alpha}^{k'k*}, \qquad (5.11б)$$

где звездочка $*$, как всегда означает величину комплексно-сопряженную. Матрица $D_{\alpha\beta}^{kk'}$, обладающая свойством (5.11б), называется эрмитовой. Известно, что собственные значения эрмитовой матрицы ω^2, следующие из уравнения (5.11а), вещественны (см. Приложение 3, п. 5). Из физических соображений ясно, что они должны быть положительны. В самом деле, отрицательные $\omega^2 = -\gamma^2 < 0$ приводят в выражении для волны (5.7) к множителю $\exp(\pm \gamma t)$, который принимает неограниченно большие значения в прошедшем или будущем, т. е. ведет к разрушению решетки. Положительность значений ω^2 может быть показана и математически из условий, налагаемых на силовые коэффициенты $\Phi_{\alpha\beta}\binom{kk'}{nn'}$, следующие из условий минимума потенциальной энергии решетки в положении равновесия.

Таким образом, в общем случае характеристическое уравнение (5.11а) степени $3s$ для ω^2 дает $3s$ различных вещественных корней $\omega_j(\boldsymbol{q})$ ($j=1, 2, 3, \ldots, 3s$), определяющих $3s$ различных ветвей колебаний. Учет симметрии кристаллической ячейки приводит к тому, что часть корней ω_j иногда совпадает, так что число различных ветвей колебаний может быть меньше $3s$. Подставляя $3s$ корней $\omega_j(\boldsymbol{q})$ в систему однородных уравнений (5.10б), получим (с точностью до постоянного множителя) $3s$ различных решений для комплексных амплитуд $A_{j\alpha}^k$.

Рассмотрим важный вопрос: в каких случаях амплитуды $A_{j\alpha}^k$ вещественны. Очевидно, что амплитуды $A_{j\alpha}^k$ вещественны (с точностью до несущественного общего множителя, который может быть и комплексным) в том случае, когда коэффициенты $D_{\alpha\beta}^{kk'}$ однородной системы (5.10б) вещественны. Для $\boldsymbol{q}=0$ и $\boldsymbol{q}\boldsymbol{a}_i = \pm\pi$ множитель $\exp[i\boldsymbol{q}(\boldsymbol{a}_{n'}-\boldsymbol{a}_n)]$ равен $+1$ или -1; но тогда из (5.10а) следует, что $D_{\alpha\beta}^{kk'}$ вещественны.

Таким образом, для бесконечно длинных и предельно коротких волн амплитуды $A_{j\alpha}^k$ в (5.7) можно считать вещественными.

Легко показать, что амплитуды $A_{j\alpha}$ можно считать вещественными для простой решетки. В этом случае решетка состоит из одинаковых атомов, расположенных в узлах $\boldsymbol{a}_n$; при этом наряду с атомом в узле $\boldsymbol{a}_n$ всегда имеется такой же атом в узле

$-\boldsymbol{a}_n$. Суммирование в (5.10а) можно провести по парам атомов $\boldsymbol{a}'_n$ и $-\boldsymbol{a}'_n$. Полагая $\boldsymbol{a}_n = 0$ и учитывая, что $\Phi_{\alpha\beta}(\boldsymbol{n}) = \Phi_{\alpha\beta}(-\boldsymbol{n})$, получим из (5.10а)

$$D_{\alpha\beta}(\boldsymbol{q}) = \frac{1}{m}\sum_{n'} \Phi_{\alpha\beta}(\boldsymbol{n}')\, e^{-i\boldsymbol{q}\boldsymbol{a}'_n} + \Phi_{\alpha\beta}(-\boldsymbol{n}')\, e^{i\boldsymbol{q}\boldsymbol{a}'_n} =$$

$$= \sum_{n'} \frac{1}{m}\Phi_{\alpha\beta}(\boldsymbol{n}')\left(e^{-i\boldsymbol{q}\boldsymbol{a}'_n} + e^{i\boldsymbol{q}\boldsymbol{a}'_n}\right) = \sum_{n'} \frac{2}{m}\Phi_{\alpha\beta}(\boldsymbol{n}')\cos \boldsymbol{q}\boldsymbol{a}'_n. \quad (5.11в)$$

Таким образом, элементы матрицы 3-го ранга $D_{\alpha\beta}(\boldsymbol{q})$ вещественны, а следовательно, могут быть выбраны вещественными и амплитуды $A_{j\alpha}$.

Из определения (5.10а) следует, что

$$D^{kk'}_{\alpha\beta}(-\boldsymbol{q}) = D^{kk'}_{\alpha\beta}(\boldsymbol{q})^*. \quad (5.12)$$

Отсюда и из эрмитовости матрицы (5.11б) следует, что характеристическое уравнение (5.11а) не меняется при замене $\boldsymbol{q}$ на $-\boldsymbol{q}$, т. е.

$$\omega_j(-\boldsymbol{q}) = \omega_j(\boldsymbol{q}). \quad (5.13)$$

Но тогда из системы (5.10б) следует, что

$$A^k_{j\alpha}(-\boldsymbol{q}) = A^k_{j\alpha}(\boldsymbol{q})^*. \quad (5.14)$$

Соотношение (5.13) является более глубоким следствием инвариантности уравнения механики при обращении времени ($t \to -t$).

3. Так как ω_j — функция $\boldsymbol{q}$, то можно в бриллюэновской зоне построить для каждой ветви колебаний j семейство поверхностей $\omega_j(\boldsymbol{q}) = \text{const}$. Структура этих *изочастотных*, или *изоэнергетических* ($\hbar\omega_j(\boldsymbol{q}) = \mathscr{E}_j(\boldsymbol{q})$) поверхностей существенно зависит от симметрии прямой решетки кристалла. В частности, симметрия в большой мере (хотя и не полностью) определяет те значения $\boldsymbol{q}$, при которых имеются касания и пересечения изоэнергетических поверхностей (разных ветвей колебаний).

Из эквивалентности волновых векторов $\boldsymbol{q}$, различающихся на $\boldsymbol{b}_g$, следует, что физические величины, зависящие от $\boldsymbol{q}$, должны быть трехмерно-периодическими функциями в $\boldsymbol{q}$-пространстве с периодами $\boldsymbol{b}_i$, где $\boldsymbol{b}_i$ — основные векторы обратной решетки. В частности, частота

$$\omega_j(\boldsymbol{q} + \boldsymbol{b}_g) = \omega_j(\boldsymbol{q}). \quad (5.15)$$

Таким образом, картина изоэнергетических поверхностей $\omega_j(\boldsymbol{q}) = \text{const}$ будет периодически повторяться в элементарных ячейках обратной решетки и зонах Бриллюэна.

Условие (5.13) показывает, что изоэнергетические поверхности $\omega_j(q) = \text{const}$, обладают в q-пространстве *центром инверсии*, или *центром симметрии*. Из условия (5.13), конечно,

не следует, что в общем случае $\omega_j(q_x) = \omega_j(-q_x)$. Для выполнения последнего необходимо, чтобы обратная решетка обладала плоскостью симметрии (отражения) $q_x = 0$, что имеет место только в том случае, если прямая решетка имеет соответствующую плоскость симметрии $x = 0$. Можно показать, что обратная решетка обладает всеми элементами симметрии прямой решетки.

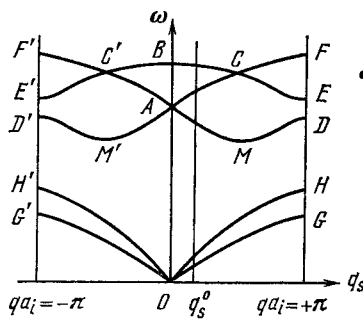


Рис. III. 9.

Выберем такое направление s в q-пространстве (это может быть, в частности, направление самого вектора q), чтобы $\omega_j(q_s) = \omega_j(-q_s)$. На рис. III. 9 изображено пять ветвей колебаний $\omega_j(q_s)$ для изменения q_s, соответствующего пределам неравенства (5.9). Мы видим, что в точках 0, A, C и C' имеет место вырождение, т. е. совпадение (пересечение) нескольких ветвей колебаний. При этом возникает вопрос, как определить заданную ветвь колебаний во всем интервале изменения q_s. В самом деле, мы можем заданную ветвь колебаний ω_j определить, например, следующими способами: $DMAC'F'$ или $DMAC'E'$, или $DMAM'D'$. Условимся поступать следующим образом: перенумеруем ветви колебаний для заданного $q_s = q_s^0$, например так:

$$\omega_1(q_s^0) < \omega_2(q_s^0) < \omega_3(q_s^0) < \omega_4(q_s^0) < \omega_5(q_s^0)$$

(см. рис. III. 9). Условимся теперь, чтобы для всех q_s выполнялось неравенство

$$\omega_1(q_s) \leqslant \omega_2(q_s) \leqslant \omega_3(q_s) \leqslant \omega_4(q_s) \leqslant \omega_5(q_s), \qquad (5.16)$$

которое определяет теперь нумерацию ветви на всем протяжении изменения q_s. Легко видеть, что в этом случае ветви $j = 1$ соответствует кривая GOG', $j = 3$—кривая $DMAM'D'$ и $j = 4$ кривая $ECAC'E'$. Только при таком выборе ветвей выполняется условие $\omega_j(q) = \omega_j(-q)$. При этом, однако, касательная $\left(\frac{\partial \omega_j}{\partial q_s}\right)_{q_s=0}$ для данной ветви испытывает, вообще говоря, скачок (см. ветви GOG' или DAD'). Например, для акустической ветви в одномерном случае мы имели соотношение (2.5)

$$\omega^2 = \omega_m^2 \sin^2 \frac{qa}{2}$$

или

$$\omega = \pm \omega_m \sin \frac{qa}{2},$$

т. е. точка $q=0$ есть точка разветвления (рис. III. 10). Отбрасывая части ветвей, соответствующие отрицательным частотам ω, мы определяем акустическую ветвь равенством

$$\omega = \omega_m \left| \sin \frac{qa}{2} \right|,$$

соответствующим кривой GOG' на рис. III. 9 и III. 10.

Конечно, нумерация ветвей в точке C (рис. III. 9) является условной, но мы ее подчиним общему правилу.

Если при $\boldsymbol{q}=0$ $\omega_j(\boldsymbol{q})$ может быть разложено в ряд по степеням $q_\alpha (\alpha = x, y, z)$ (это имеет место, когда в этой точке нет вырождения), то

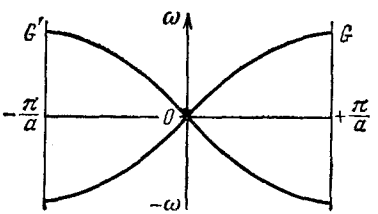


Рис. III.10.

$$\omega_j(\boldsymbol{q}) = \omega_j(0) + \sum_\alpha r_\alpha q_\alpha + \sum_{\alpha\beta} R_{\alpha\beta} q_\alpha q_\beta, \tag{5.17}$$

где

$$r_\alpha = \left(\frac{\partial \omega_j}{\partial q_\alpha}\right)_0, \quad R_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial^2 \omega_j}{\partial q_\alpha \partial q_\beta}\right)_{00}.$$

Заменяя $\boldsymbol{q}$ на $-\boldsymbol{q}$ и используя соотношение (5.13), получим $\sum_\alpha r_\alpha q_\alpha = 0$, что ввиду произвольности q_α дает

$$r_\alpha = \left(\frac{\partial \omega_j}{\partial q_\alpha}\right) = 0.$$

Приводя тензор $R_{\alpha\beta}$ к главным осям, получим

$$\omega_j(\boldsymbol{q}) = \omega_j^0 + \sum_\alpha R_\alpha q_\alpha^2, \tag{5.18}$$

т. е., в отсутствие вырождения в точке $q=0$, разложение $\omega_j - \omega_j^0$ начинается с квадратичных членов. Этот случай изображен на рис. III. 9 точкой B.

Для кубического кристалла компоненты тензора R_α вырождаются в скаляр R, так что

$$\omega_j(\boldsymbol{q}) = \omega_j^0 + Rq^2, \tag{5.18а}$$

т. е. изоэнергетические поверхности превращаются в сферы.

В точке A разложение ω_j в степенной ряд невозможно (производная ω_j неоднозначна), поэтому здесь, вообще говоря, для каждой ветви

$$\left(\frac{\partial \omega_j}{\partial q}\right)_0 \neq 0.$$

Наконец, отметим, что в общем случае экстремумы $\omega_j(q_s)$ могут иметь место как в центре и на границах бриллюэновской

зоны, так и в некоторых точках внутри нее (например, в точках M и M' на рис. III. 9).

5. Рассмотрим характер решений уравнений (5.10) в предельном случае очень длинных волн ($\boldsymbol{q} \to 0$), когда, как мы указывали выше, амплитуды $A^k_\alpha(0)$ вещественны и, следовательно, характеризуют реальные отклонения атомов $u^k_{n\alpha}$ от положений равновесия.

Из (5.10) и (5.10а) следует, что при $\boldsymbol{q}=0$

$$\omega^2(0)\, A^k_\alpha(0) = \sum_{k'\beta n'} \frac{1}{\sqrt{m_k m_{k'}}} \Phi_{\alpha\beta}\begin{pmatrix} k & k' \\ n & n' \end{pmatrix} A^{k'}_\beta(0). \tag{5.19}$$

Положим для j-й ветви амплитуды колебаний (см. (5.7))

$$A^{k'}_{j\beta}(0)/\sqrt{m_{k'}} = \mathrm{const}_{(k')}, \tag{5.20}$$

т. е. не зависит от номера атома k'. В этом случае правая часть (5.19) равна

$$\frac{1}{\sqrt{m_k}} \sum_\beta \frac{A^{k'}_{j\beta}(0)}{\sqrt{m_{k'}}} \sum_{k'n'} \Phi_{\alpha\beta}\begin{pmatrix} kk' \\ nn' \end{pmatrix} = 0$$

в силу (5.3б); но тогда из (5.10) следует, что

$$\omega_j(0) = 0 \tag{5.21}$$

(так как нас не интересует случай $A^k_{j\alpha}(0)=0$ для всех трех значений $\alpha = x, y, z$). С другой стороны, для того чтобы имело место (5.21), достаточно, чтобы $A^k_{j\alpha}(0) \neq 0$ для одного α, поэтому естественно предположить, что существуют три ветви колебаний ($j=1, 2, 3$), для которых $\omega_j(0)=0$, т. е. частота стремится к нулю, когда $\boldsymbol{q} \to 0$. Эти три ветви колебаний, для которых характер движения атомов имеет вид (5.20), называются *акустическими*, и они аналогичны акустической ветви, рассмотренной в одномерном случае (§ 3).

Рассмотрим теперь случай, когда $\boldsymbol{q}=0$, но не выполняется (5.20) и следующее из него условие (5.21).

Перепишем (5.19) в виде

$$\omega^2(0)\, m_k \frac{A^k_\alpha(0)}{\sqrt{m_k}} = \sum_{k'\beta n'} \frac{1}{\sqrt{m_{k'}}} \Phi_{\alpha\beta}\begin{pmatrix} k & k' \\ n & n' \end{pmatrix} A^{k'}_\beta(0)$$

и просуммируем обе части по k:

$$\omega^2(0) \sum_k m_k \frac{A^k_\alpha(0)}{\sqrt{m_k}} = \sum_{k'\beta} \frac{A^{k'}_\beta(0)}{\sqrt{m_{k'}}} \sum_{n'k} \Phi_{\alpha\beta}\begin{pmatrix} k & k' \\ n & n' \end{pmatrix}.$$

Последняя сумма в правой части, согласно (5.3б), равна нулю. Таким образом, левая часть равенства равна нулю, а так как мы предположили, что $\omega^2(0) \neq 0$, то сумма по k в левой части

равна нулю. Но тогда, используя (5.7), получим

$$\sum_k m_k u_{n\alpha}^k(0) = 0, \tag{5.22}$$

т. е. движение атомов происходит при этом так, что центр тяжести ячейки остается на месте. Такие ветви колебаний получили название *оптических*; их аналог в одномерном случае был рассмотрен в § 3.

Мы видим, что наряду с тремя акустическими должны существовать $3s-3$ оптические ветви, для которых $\omega_j(0) \neq 0$ $(j=4, 5, \ldots, 3s)$.

Можно показать, так же как в одномерном случае, что при $\boldsymbol{q}\boldsymbol{a}_i \ll 1$ частоты акустических ветвей $\omega_j = v_{0j} q$ $(j=1, 2, 3)$, где скорость волн v_{0j} зависит не только от ветви j, но и от направления волнового вектора $\boldsymbol{q}$. Для оптических ветвей зависимость ω_j от $\boldsymbol{q}$ вблизи нуля в отсутствие вырождения определяется выражением (5.18), т. е. имеет экстремум (обычно максимум). На гранях зоны Бриллюэна в отсутствие вырождения ветвей колебаний частоты $\omega_j(\boldsymbol{q})$ также экстремальны. На рис. III.11 схематически представлены акустические и оптические ветви (в отсутствие вырождения) для некоторого направления в кристалле, имеющего два атома в элементарной ячейке.

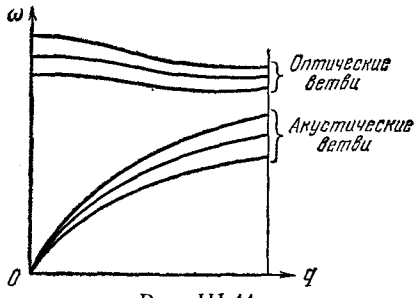


Рис. III.11.

Для того чтобы иметь полную картину колебаний атомов в кристалле, необходимо решить уравнение (5.11) для всех значений $\boldsymbol{q}$ в интервале $-\pi \leqslant \boldsymbol{q}\boldsymbol{a}_i \leqslant +\pi$, или для всех точек бриллюэновской зоны. Если даже известны силовые коэффициенты $\Phi_{\alpha\beta}\begin{pmatrix} k & k' \\ n & n' \end{pmatrix}$, то это можно выполнить только путем численных расчетов. Такие расчеты показывают, что в трехмерном кристалле акустические и оптические ветви приблизительно имеют вид, подобный изображенному на рис. III. 11.

6. Совершенно аналогично тому, как это имело место в одномерном случае (§ 2 п. 3), заменим граничные условия на поверхности основной области кристалла условиями цикличности Борна—Кармана. В случае трехмерного кристалла они не имеют столь наглядного смысла, как в случае одномерного, когда мы могли всю цепочку атомов расположить по окружности, сомкнув 1-й и G-й атомы. Однако может быть показано[1]), что в трех-

[1]) Борн М., Кунь Х., Динамическая теория кристаллических решеток.— М.: ИЛ, 1958, с. 453.

мерном кристалле для определения его объемных свойств условия цикличности эквивалентны граничным условиям на свободной поверхности основной области кристалла.

В трехмерном кристалле условия цикличности Борна—Кармана сводятся к требованию, чтобы волновое поле $u^k_{n\alpha}$, пропорциональное e^{iqa_n}, оставалось неизменным при смещении на любой из векторов Ga_i $(i=1, 2, 3)$. Для этого необходимо, чтобы $e^{iqGa_i}=1$, т. е. $G(\boldsymbol{qa}_i)=2\pi g_i$, где g_i — целое число. Таким образом,

$$\varphi_i \equiv (\boldsymbol{qa}_i) = \frac{2\pi}{G} g_i \quad (i=1, 2, 3). \tag{5.23}$$

Сравнивая это соотношение с (I.3.7), мы видим, что (5.23) выполняется, если

$$\boldsymbol{q} = \frac{1}{G}\boldsymbol{b}_g = \frac{1}{G}(g_1\boldsymbol{b}_1 + g_2\boldsymbol{b}_2 + g_3\boldsymbol{b}_3), \tag{5.24}$$

где $\boldsymbol{b}_i$ — основные векторы обратной решетки. Это представление вектора $\boldsymbol{q}$, аналогичное (2.8), носит формальный характер, поскольку G — произвольное большое число (удобно считать его нечетным), от которого не должны зависеть физически наблюдаемые величины.

Используя неравенство (5.9), получим

$$-G/2 < g_i < G/2 \tag{5.25}$$

аналогично (2.8а). Мы видим, что проекция $\boldsymbol{q}$ на каждый из векторов $\boldsymbol{a}_i$ может принимать G квазидискретных значений, и, следовательно, $\boldsymbol{q}$ имеет всего $G^3=N$ различных значений. Поставим теперь вопрос о числе различных колебаний (возможных значений $\boldsymbol{q}$) для одной ветви колебаний в объеме $d\tau_q = dq_x\, dq_y\, dq_z$ пространства волнового вектора. Очевидно, это число

$$dz = dg_1\, dg_2\, dg_3 = \left(\frac{G}{2\pi}\right)^3 d\varphi_1\, d\varphi_2\, d\varphi_3, \tag{5.26}$$

где

$$\varphi_i = q_x a_{ix} + q_y a_{iy} + q_z a_{iz}.$$

Перейдем от переменных φ_i к переменным q_x, q_y, q_z. Для этого вычислим якобиан преобразования [1]):

$$\begin{vmatrix} \frac{\partial\varphi_1}{\partial q_x} & \frac{\partial\varphi_2}{\partial q_x} & \frac{\partial\varphi_3}{\partial q_x} \\ \frac{\partial\varphi_1}{\partial q_y} & \frac{\partial\varphi_2}{\partial q_y} & \frac{\partial\varphi_3}{\partial q_y} \\ \frac{\partial\varphi_1}{\partial q_z} & \frac{\partial\varphi_2}{\partial q_z} & \frac{\partial\varphi_3}{\partial q_z} \end{vmatrix} = \begin{vmatrix} a_{1x} & a_{2x} & a_{3x} \\ a_{1y} & a_{2y} & a_{3y} \\ a_{1z} & a_{2z} & a_{3z} \end{vmatrix} = (\boldsymbol{a}_1[\boldsymbol{a}_2\boldsymbol{a}_3]) = \Omega_0.$$

[1]) В. И. Смирнов. Курс высшей математики.— 21 изд.— М.: Наука, 1974, т. 2, с. 195.

Таким образом,

$$dz = \left(\frac{G}{2\pi}\right)^3 d\varphi_1\, d\varphi_2\, d\varphi_3 = \left(\frac{G}{2\pi}\right)^3 \Omega_0\, dq_x\, dq_y\, dq_z = \frac{V}{(2\pi)^3}\, d\tau_q\,. \qquad (5.27)$$

Для одной ветви колебаний кристалла полное число колебаний

$$z = \int dz = \frac{V}{(2\pi)^3} \iiint d\tau_q = \frac{V}{(2\pi)^3}\,\frac{(2\pi)^3}{\Omega_0} = N.$$

Для всех ветвей колебаний сложного кристалла это число всегда равно $3sN$, т. е. числу степеней свободы атомов основной области.

Во многих случаях, например при изучении статистических вопросов, требуется знать число колебаний для j-й ветви в интервале от ω до $\omega + d\omega$. Рассмотрим в q-пространстве поверхности $\omega_j(\boldsymbol{q}) = \text{const}$ и положим

$$d\tau_q = d\sigma\, dq_\perp, \qquad (5.28)$$

где $d\sigma$ — элемент этой поверхности, а $dq_\perp$ — бесконечно малое приращение по нормали к $d\sigma$.

Очевидно, что

$$d\omega = |\nabla_q \omega_j|\, dq_\perp, \qquad (5.29)$$

где ∇_q — градиент в $\boldsymbol{q}$-пространстве. Подставляя в (5.27) (5.28) и (5.29) и интегрируя по поверхности постоянной частоты, получим для числа колебаний в j-й ветви, в интервале частот $(\omega, \omega + d\omega)$, выражение

$$dz_\omega = \frac{V}{(2\pi)^3} \int \frac{d\sigma}{|\nabla_q \omega_j|}\, d\omega = V g_j(\omega)\, d\omega, \qquad (5.30)$$

где

$$g_j(\omega) = \frac{1}{(2\pi)^3} \int \frac{d\sigma}{|\nabla_q \omega_j|} \qquad (5.31)$$

— *функция распределения частот* или *собственных колебаний для ветви* j. Мы видим, что для ее определения необходимо знать зависимость $\omega_j = \omega_j(\boldsymbol{q})$, т. е. *закон дисперсии*.

Если нас интересует полное число колебаний во всех ветвях дисперсии, то необходимо пользоваться полной функцией распределения частот

$$g(\omega) = \sum_{j=1}^{3s} g_j(\omega) = \frac{1}{(2\pi)^3} \sum_{j=1}^{3s} \int \frac{d\sigma}{|\nabla_q \omega_j|}\,. \qquad (5.32)$$

Простой пример функции распределения частот был дан для одномерной решетки в (2.10).

В некоторых случаях удобнее пользоваться другой функцией распределения частот $\tilde{g}(\omega^2)$, имеющей тот смысл, что $\tilde{g}(\omega^2)\, d\omega^2$ —

число колебаний в интервале от ω^2 до $\omega^2+d\omega^2$. Очевидно, что

$$g(\omega)=2\omega\tilde{g}(\omega^2), \tag{5.33}$$

так как $\tilde{g}(\omega^2)\,d\omega^2=\tilde{g}(\omega^2)\,2\omega\,d\omega$.

§ 6. Нормальные координаты колебаний кристаллической решетки

1. В § 4 мы рассмотрели нормальные координаты для простой одномерной решетки.

При этом мы вначале ввели комплексные нормальные координаты a_q, а затем перешли к вещественным нормальным координатам x_q и соответствующим им обобщенным импульсам p_q.

Отличительной особенностью нормальных координат является то, что каждая из них гармонически зависит от времени, т. е. пропорциональна $\exp(i\omega_q t)$. Это эквивалентно утверждению, что функция Гамильтона состоит из суммы квадратов (с соответствующими коэффициентами) координат и импульсов, т. е. имеет вид (4.11).

Очевидно, что наиболее общее движение атомов сложной трехмерной кристаллической решетки состоит из суммы выражений (5.7) с различными волновыми векторами $\boldsymbol{q}$ и частотами $\omega(\boldsymbol{q})$. В самом деле, такая сумма тоже будет удовлетворять уравнениям движения (5.6) в силу их линейности.

Для того чтобы ввести нормальные координаты для трехмерной решетки, определим *собственные векторы* $e_{jk}(\boldsymbol{q})$ (в общем случае комплексные) динамические матрицы $D^{kk'}_{\alpha\beta}(\boldsymbol{q})$ (см. (5.10а)); имеем (см. Приложение 3, п. 5)

$$\sum_{k'\beta} D^{kk'}_{\alpha\beta}(\boldsymbol{q})\,e_{jk'\beta}(\boldsymbol{q})=\omega_j^2(\boldsymbol{q})\,e_{jk\alpha}(\boldsymbol{q}), \tag{6.1}$$

где $e_{jk\alpha}$ $(\alpha=x,\ y,\ z)$ — составляющие собственного вектора, а $\omega_j^2(\boldsymbol{q})$ — j-е собственное значение динамической матрицы. Так как динамическая матрица эрмитова, то ее собственные векторы ортогональны (см. Приложение 3, п. 5), и так как они определяются из однородной системы (6.1) с точностью до множителя,— они могут быть нормированы на единицу. Таким образом,

$$\sum_{\alpha k} e_{jk\alpha}e^*_{j'k\alpha}=\delta_{jj'} \tag{6.2}$$

и

$$\sum_{j} e_{jk\alpha}e^*_{jk'\beta}=\delta_{kk'}\delta_{\alpha\beta}. \tag{6.2а}$$

Собственные векторы $e_{jk\alpha}(\boldsymbol{q})$ отличаются от комплексных амплитуд $A^k_{j\alpha}(\boldsymbol{q})$, удовлетворяющих уравнениям (5.10), только условием нормировки.

Из условий (5.12) следует (5.14), т. е.

$$e_{jk\alpha}(\boldsymbol{q}) = e^*_{jk\alpha}(-\boldsymbol{q}), \tag{6.3}$$

так как при замене $\boldsymbol{q}$ на $-\boldsymbol{q}$ все коэффициенты уравнений (6.1) меняются на комплексно-сопряженные.

2. Введем комплексные нормальные координаты $a_j(\boldsymbol{q}, t)$ посредством соотношения

$$u^k_{n\alpha} = \frac{1}{\sqrt{Nm_k}} \sum_{qj} e_{jk\alpha}(\boldsymbol{q})\, a_j(\boldsymbol{q}, t)\, e^{iqa_n}, \tag{6.4}$$

где $u^k_{n\alpha}$—α-я составляющая отклонения атома $(\boldsymbol{n}, k)$ от положения равновесия, N—число элементарных ячеек основной области кристалла, m_k—масса атома k-го сорта.

Легко показать, что

$$a_j(\boldsymbol{q}) = a^*_j(-\boldsymbol{q}). \tag{6.5}$$

В самом деле, только в этом случае отклонения $u^k_{n\alpha}$ вещественны; действительно, из (6.4), (6.3) и (6.5) следует

$$(u^k_{n\alpha})^* = \frac{1}{\sqrt{Nm_k}} \sum_{qj} e^*_{jk\alpha}(\boldsymbol{q})\, a^*_j(\boldsymbol{q})\, e^{-iqa_n} =$$

$$= \frac{1}{\sqrt{Nm_k}} \sum_{qj} e_{jk\alpha}(-\boldsymbol{q})\, a_j(-\boldsymbol{q})\, e^{-iqa_n} = u^k_{n\alpha},$$

так как суммирование по $-\boldsymbol{q}$ эквивалентно суммированию по $\boldsymbol{q}$. Учитывая, что $\boldsymbol{q}$ пробегает N квазидискретных значений, а j—$3s$ значений, то имеется всего $3sN$ комплексных величин $a_j(\boldsymbol{q})$. Из-за условия (6.5) это эквивалентно заданию $3sN$ вещественных координат ($3sN$ равно числу степеней свободы системы).

Выражение (6.4) для смещений атомов может быть записано в несколько иной для некоторых целей более удобной форме.

Проведем через начало координат $\boldsymbol{q}$-пространства плоскость; тогда векторы $\boldsymbol{q}$ и $-\boldsymbol{q}$ расположены по разные стороны от этой плоскости. Запишем выражение (6.4) через суммы по $\boldsymbol{q}$, взятые по разные стороны от плоскости:

$$u^k_{n\alpha} = \frac{1}{\sqrt{Nm_k}} \left\{ \sum_{q>0,\ j} e_{jk\alpha}(\boldsymbol{q})\, a_j(\boldsymbol{q}, t)\, e^{iqa_n} + \right.$$

$$\left. + \sum_{q<0,\ j} e_{jk\alpha}(\boldsymbol{q})\, a_j(\boldsymbol{q}, t)\, e^{iqa_n} \right\}.$$

Меняя во второй сумме $\boldsymbol{q}$ на $-\boldsymbol{q}$ и используя (6.3) и (6.5), получим

$$u^k_{n\alpha} = \frac{1}{\sqrt{Nm_k}} \sum_{qj}{}' \left\{ e_{jk\alpha}(\boldsymbol{q})\, a_j(\boldsymbol{q}, t)\, e^{iqa_n} + e^*_{jk\alpha}(\boldsymbol{q})\, a^*_j(\boldsymbol{q}, t)\, e^{-iqa_n} \right\}, \tag{6.4а}$$

где штрих у суммы $\sum'$ означает, что суммирование по $\boldsymbol{q}$ ведется по половине зоны Бриллюэна.

Для того чтобы выразить комплексные координаты $a_j(\boldsymbol{q}, t)$ через вещественные смещения $u^k_{n\alpha}$, умножим обе части равенства (6.4) на $\sqrt{m_k}e^*_{j'k\alpha}(\boldsymbol{q}')\exp(-i\boldsymbol{q}'\boldsymbol{a}_n)$ и просуммируем их по n, k, α. Суммирование в правой части по $\boldsymbol{n}$ множителя $\exp[i(\boldsymbol{q}-\boldsymbol{q}')\boldsymbol{a}_n]$ дает, согласно (П. 6.4), величину $N\delta_{qq'}$; суммирование по k, α множителя $e_{jk\alpha}(\boldsymbol{q})e^*_{j'k\alpha}$ дает, согласно (6.2), величину $\delta_{jj'}$. Окончательно после суммирования по $\boldsymbol{q}$ и j получим для правой части $\sqrt{N}a_{j'}(\boldsymbol{q}')$. Переобозначая j' и $\boldsymbol{q}'$ на j и $\boldsymbol{q}$, получим

$$a_j(\boldsymbol{q}) = \frac{1}{\sqrt{N}}\sum_{nk\alpha}\sqrt{m_k}u^k_{n\alpha}e_{jk\alpha}(\boldsymbol{q})e^{-i\boldsymbol{q}\boldsymbol{a}_n}. \tag{6.6}$$

Для того чтобы убедиться в том, что $a_j(\boldsymbol{q})$ — нормальные координаты системы, выразим через них кинетическую и потенциальную энергии кристалла.

Кинетическая энергия атомов, в расчете на основную область кристалла, равна

$$\mathscr{T} = \frac{1}{2}\sum_{nk\alpha}m_k(\dot{u}^k_{n\alpha})^2 = \frac{1}{2N}\sum_{nk\alpha}\left[\sum_{qj}e_{jk\alpha}(\boldsymbol{q})\dot{a}_j(\boldsymbol{q}, t)e^{i\boldsymbol{q}\boldsymbol{a}_n}\right]^2, \tag{6.7}$$

где мы воспользовались соотношением (6.4).

Так как квадратная скобка правой части (6.7) вещественна, то ее квадрат может быть формально представлен как произведение скобки на комплексно сопряженную. Таким образом,

$$\mathscr{T} = \frac{1}{2N}\sum_{nk\alpha}\sum_{\boldsymbol{q}j\boldsymbol{q}'j'}[e_{jk\alpha}(\boldsymbol{q})\dot{a}_j(\boldsymbol{q}, t)e^{i\boldsymbol{q}\boldsymbol{a}_n}e^*_{j'k\alpha}(\boldsymbol{q}')\dot{a}^*_{j'}(\boldsymbol{q}', t)e^{-i\boldsymbol{q}'\boldsymbol{a}_n}]. \tag{6.8}$$

Из (П. 6.4) следует, что

$$\sum_n e^{i(\boldsymbol{q}-\boldsymbol{q}')\boldsymbol{a}_n} = N\delta_{\boldsymbol{q}\boldsymbol{q}'},$$

поэтому суммирование по $\boldsymbol{q}'$ позволяет во всех множителях положить $\boldsymbol{q}' = \boldsymbol{q}$; после этого суммирование по k, α произведения $e_{jk\alpha}e^*_{j'k\alpha}$ дает множитель $\delta_{jj'}$ (см. (6.2)). Наконец, суммируя по j', получим из (6.8)

$$\mathscr{T} = \frac{1}{2}\sum_{qj}|\dot{a}_j(\boldsymbol{q}, t)|^2. \tag{6.9}$$

Потенциальная энергия основной области кристалла в квадратичном приближении по смещениям атомов, согласно (5.2), равна

$$\Phi = \frac{1}{2}\sum_{nn'k\alpha k'\beta}\Phi_{\alpha\beta}\begin{pmatrix}kk'\\nn'\end{pmatrix}u^k_{n\alpha}u^{k'}_{n'\beta} =$$

$$= \frac{1}{2N}\sum_{nn'k\alpha k'\beta}\Phi_{\alpha\beta}\begin{pmatrix}kk'\\nn'\end{pmatrix}\frac{1}{\sqrt{m_k}}\sum_{q,\,j}e^*_{jk\alpha}(\boldsymbol{q})a^*_j(\boldsymbol{q})e^{-i\boldsymbol{q}\boldsymbol{a}_n}\times$$

$$\times\frac{1}{\sqrt{m_{k'}}}\sum_{\boldsymbol{q}'j'}e_{j'k'\beta}(\boldsymbol{q}')a_{j'}(\boldsymbol{q}')e^{i\boldsymbol{q}'\boldsymbol{a}'_n}, \tag{6.10}$$

где мы выразили смещения атомов через (6.4) и для удобства вычислений записали вещественную величину $u^k_{n\alpha}$ через $(u^k_{n\alpha})^*$.

Преобразуем сумму по $\boldsymbol{n}$ и $\boldsymbol{n'}$:

$$\sum_{nn'} \frac{1}{\sqrt{m_k m_{k'}}} \Phi_{\alpha\beta}\begin{pmatrix} kk' \\ nn' \end{pmatrix} e^{i(q'a'_n - qa_n)} =$$

$$= \sum_n e^{i(q'-q)a_n} \sum_{n'} \frac{1}{\sqrt{m_k m_{k'}}} \Phi_{\alpha\beta}\begin{pmatrix} kk' \\ nn' \end{pmatrix} e^{iq'(a'_n - a_n)} = N\delta_{qq'} D^{kk'}_{\alpha\beta}(q'),$$

где мы использовали (П. 6.4) и (5.10а)[1]. Подставляя это выражение в (6.10), суммируя по q' и используя (6.1), получим

$$\Phi = \frac{1}{2} \sum_{jk\alpha j'k'\beta qq'} D^{kk'}_{\alpha\beta}(q')\, \delta_{qq'}\, e^*_{jk\alpha}(q)\, a^*_j(q)\, e_{j'k'\beta}(q')\, a_{j'}(q') =$$

$$= \frac{1}{2} \sum_{k\alpha jj'q} \omega^2_{j'}(q)\, e_{j'k\alpha}(q)\, e^*_{jk\alpha}(q)\, a_{j'}(q)\, a^*_j(q).$$

Суммируя по k, α с использованием (6.2) и суммируя по j', получим окончательно

$$\Phi = \frac{1}{2} \sum_{qj} \omega^2_j(q)\, |a_j(q)|^2. \tag{6.11}$$

Таким образом, для полной энергии основной области кристалла получим

$$\mathscr{E} = \mathscr{T} + \Phi = \frac{1}{2} \sum_{qj} [|\dot{a}_j(q)|^2 + \omega^2_j(q)\, |a_j(q)|^2]. \tag{6.12}$$

Из структуры этого выражения видно, что энергия колебаний атомов кристалла равна сумме энергий колебаний отдельных независимых «осцилляторов» с «кинетической энергией» $^1/_2 |\dot{a}_j(q)|^2$ и «потенциальной энергией» $^1/_2 \omega^2_j(q) |a_j(q)|^2$; поэтому величины $a_j(q)$ уместно назвать *комплексными нормальными координатами*.

3. Так как законы механики сформулированы для вещественных координат и скоростей, то желательно в выражении (6.12) перейти к вещественным нормальным координатам. Наиболее непосредственно их можно ввести, полагая

$$a_j(q) = \frac{1}{\sqrt{2}} [Q_{1j}(q) + iQ_{2j}(q)], \tag{6.13}$$

где $Q_{1j}(q)$ и $Q_{2j}(q)$ вещественны. На первый взгляд кажется, что число их в два раза больше числа степеней свободы, однако из (6.5) следует, что

$$Q_{1j}(q) = Q_{1j}(-q), \quad Q_{2j}(q) = -Q_{2j}(-q). \tag{6.14}$$

[1]) Надо учесть, что $\sum_{n'}(\ldots)$ не зависит от $\boldsymbol{n}$, так как выражение, стоящее под знаком суммы, зависит только от разности $\boldsymbol{n'-n}$.

В силу этого число независимых вещественных нормальных координат равно $3sN$, т. е. числу степеней свободы (как уже указывалось после равенства (6.5)).

Для того чтобы выразить энергию $\mathscr{E}$ (6.12) через независимые координаты Q_{1j} и Q_{2j}, мы проведем в $\boldsymbol{q}$-пространстве плоскость через начало координат и после подстановки (6.13) в (6.12) будем суммировать по $\boldsymbol{q}$ только по одной стороне от плоскости, т. е. по половине бриллюэновской зоны; тогда

$$\mathscr{E} = \frac{1}{4}\sum_{qj}{}' \{[\dot{Q}_{1j}^2(\boldsymbol{q}) + \omega_j^2(\boldsymbol{q})\, Q_{1j}^2(\boldsymbol{q})] + [\dot{Q}_{2j}^2(\boldsymbol{q}) + \omega_j^2(\boldsymbol{q})\, Q_{2j}^2(\boldsymbol{q})]\}. \quad (6.15)$$

Можно показать, что вещественным нормальным координатам (6.13) соответствуют стоячие волны в кристалле, смещенные друг относительно друга на 1/4 длины волны[1]). Для приложений более удобными являются бегущие волны. Вещественные нормальные координаты и скорости (импульсы) для этого случая были впервые введены Р. Пайерлсом посредством *канонического преобразования*[2])

$$a_j(\boldsymbol{q}) = \tilde{a}_j(\boldsymbol{q}) + \tilde{a}_j^*(-\boldsymbol{q}). \quad (6.16)$$

Здесь

$$\tilde{a}_j(\boldsymbol{q}) = \frac{1}{2}\left[Q_j(\boldsymbol{q}) + \frac{i}{\omega_j(\boldsymbol{q})}\dot{Q}_j(\boldsymbol{q})\right], \quad (6.16а)$$

где $Q_j(\boldsymbol{q})$ — вещественные нормальные координаты. Очевидно, что условие (6.5) выполняется автоматически. Ниже будет показано, что $Q_j(\pm\boldsymbol{q})$ — действительно нормальные координаты, и поэтому

$$\ddot{Q}_j(\pm\boldsymbol{q}) = -\omega_j^2(\boldsymbol{q})\, Q_j(\pm\boldsymbol{q}), \quad (6.17)$$

где учтено, что $\omega_j(-\boldsymbol{q}) = \omega_j(\boldsymbol{q})$ (см. (5.13)).

Из (6.16) и (6.17) получим

$$\dot{a}_j(\boldsymbol{q}) = {}^1/_2[\dot{Q}_j(\boldsymbol{q}) - i\omega_j(\boldsymbol{q})\, Q_j(\boldsymbol{q})] + {}^1/_2[\dot{Q}_j(-\boldsymbol{q}) + i\omega_j(\boldsymbol{q})\, Q_j(-\boldsymbol{q})]. \quad (6.18)$$

Подставляя теперь (6.16) и (6.18) в (6.12), получим

$$\mathscr{E} = \frac{1}{4}\sum_{qj}\{[\dot{Q}_j^2(\boldsymbol{q}) + \omega_j^2(\boldsymbol{q})\, Q_j^2(\boldsymbol{q})] + [\dot{Q}_j^2(-\boldsymbol{q}) + \omega_j^2(\boldsymbol{q})\, Q_j^2(-\boldsymbol{q})]\} = \frac{1}{2}\sum_{qj}[\dot{Q}_j^2(\boldsymbol{q}) + \omega_j^2(\boldsymbol{q})\, Q_j^2(\boldsymbol{q})], \quad (6.19)$$

так как суммирования по $-\boldsymbol{q}$ и $\boldsymbol{q}$ эквивалентны, а из (5.13)

---

[1]) Борн М., Кунь Х. Динамическая теория кристаллических решеток.— М.: ИЛ, 1958, § 38.

[2]) Ландау Л. Д., Лифшиц Е. М. Механика.— 4 изд.— М.: Наука, 1973, § 45.

$\omega_j(-\boldsymbol{q}) = \omega_j(\boldsymbol{q})$. Из (6.19) видно, что $Q_j(\boldsymbol{q})$ действительно являются нормальными координатами, что оправдывает выражение (6.17) и подтверждает, что преобразование (6.16) — каноническое.

Вводя вещественные обобщенные импульсы

$$P_j(\boldsymbol{q}) = \frac{\partial \mathscr{E}}{\partial \dot{Q}_j} = \dot{Q}_j(\boldsymbol{q}), \tag{6.20}$$

сопряженные координатам $Q_j(\boldsymbol{q})$, можно записать (6.19) в виде

$$\mathscr{E} = \sum_{qj} \left[\frac{1}{2} P_j^2(\boldsymbol{q}) + \frac{1}{2} \omega_j^2(\boldsymbol{q}) Q_j^2(\boldsymbol{q}) \right] = \mathscr{H}(Q, P), \tag{6.21}$$

где $\mathscr{H}(Q, P)$ — функция Гамильтона системы.

Подставляя (6.16) в (6.4), получим для смещений атомов

$$u_{n\alpha}^k = \frac{1}{\sqrt{Nm_k}} \sum_{qj} e_{jk\alpha}(\boldsymbol{q}) [\tilde{a}_j(\boldsymbol{q}) + \tilde{a}_j^*(-\boldsymbol{q})] e^{i\boldsymbol{q}\boldsymbol{a}_n}. \tag{6.22}$$

Если во втором слагаемом, содержащем множитель $\tilde{a}_j^*(-\boldsymbol{q})$, перейти при суммировании от $\boldsymbol{q}$ к $-\boldsymbol{q}$ (что, очевидно, равносильно) и воспользоваться явным видом (6.16), (6.20) и (6.3), то

$$u_{n\alpha}^k = \frac{1}{\sqrt{Nm_k}} \sum_{qj} \operatorname{Re} \left\{ e_{j\alpha k}(\boldsymbol{q}) \left[Q_j(\boldsymbol{q}) + \frac{i}{\omega_j(\boldsymbol{q})} P_j(\boldsymbol{q}) \right] e^{i\boldsymbol{q}\boldsymbol{a}_n} \right\} \tag{6.23}$$

($\operatorname{Re}\{\ldots\}$ — вещественная часть выражения, в фигурной скобке).

Покажем, что возбуждению одной нормальной координаты $Q_j(\boldsymbol{q})$ соответствует бегущая волна с волновым вектором $\boldsymbol{q}$ и частотой $\omega_j(\boldsymbol{q})$.

Если только одна нормальная координата $Q_j(\boldsymbol{q})$ отлична от нуля, то сумма в (6.22) превращается в одно слагаемое

$$u_{n\alpha}^k = \frac{1}{\sqrt{Nm_k}} \operatorname{Re} \left\{ e_{j\alpha k}(\boldsymbol{q}) \left[Q_j(\boldsymbol{q}) + \frac{i}{\omega_j(\boldsymbol{q})} P_j(\boldsymbol{q}) \right] e^{i\boldsymbol{q}\boldsymbol{a}_n} \right\}. \tag{6.24}$$

Так как $Q_j(\boldsymbol{q})$ — нормальная координата, то $Q_j(\boldsymbol{q})$ и $\frac{i}{\omega_j(\boldsymbol{q})} P_j(\boldsymbol{q})$ гармонически зависят от времени, т. е. пропорциональны $\exp(-i\omega_j(\boldsymbol{q})t)$. Обозначая комплексную амплитуду в фигурной скобке через $C_\alpha(\boldsymbol{q}) e^{i\varphi}$ (множитель, стоящий перед $\operatorname{Re}\{\ldots\}$, включен в $C_\alpha(\boldsymbol{q})$), получим

$$u_{n\alpha}^k = \operatorname{Re} \{ C_\alpha(\boldsymbol{q}) e^{i(\boldsymbol{q}\boldsymbol{a}_n) - \omega_j(\boldsymbol{q})t + \varphi} \}, \tag{6.25}$$

откуда следует

$$u_{n\alpha}^k = C_\alpha(\boldsymbol{q}) \cos(\boldsymbol{q}\boldsymbol{a}_n - \omega_j(\boldsymbol{q})t + \varphi), \tag{6.26}$$

где $\boldsymbol{C}(\boldsymbol{q})$ — амплитуда плоской волны, распространяющейся в направлении волнового вектора $\boldsymbol{q}$, с частотой $\omega_j(\boldsymbol{q})$; φ — фаза этой волны. Мы видим, что нормальным координатам $Q_j(\boldsymbol{q})$, введенным посредством канонического преобразования (6.16), соответствуют бегущие волны (6.26).

4. В предыдущем параграфе было показано, что для простой решетки (решетки Браве) амплитуды $A_{j\alpha}$ вещественны. Очевидно, то же имеет место для собственных векторов $e_{j\alpha}$[1]).

Из (6.4а) следует, что для простой решетки

$$u_{n\alpha}=\frac{1}{\sqrt{NM}}\sum_{qj}{}' e_{j\alpha}(\boldsymbol{q})\{a_j(\boldsymbol{q},\ t)\,e^{i\boldsymbol{q}\boldsymbol{a}_n}+a_j^*(\boldsymbol{q},\ t)\,e^{-i\boldsymbol{q}\boldsymbol{a}_n}\}. \quad (6.27)$$

Здесь M — масса атома, а штрих у суммы означает, что суммирование по $\boldsymbol{q}$ ведется по половине бриллюэновской зоны.

Наконец, из (6.23) для простой решетки следует

$$u_{n\alpha}=\frac{1}{\sqrt{NM}}\sum_{qj}e_{j\alpha}(\boldsymbol{q})\left[Q_j(\boldsymbol{q})\cos\boldsymbol{q}\boldsymbol{a}_n-\frac{P_j(\boldsymbol{q})}{\omega_j(\boldsymbol{q})}\sin\boldsymbol{q}\boldsymbol{a}_n\right]. \quad (6.28)$$

§ 7. Колебания простой кубической решетки

1. Рассмотрим простую кубическую решетку, в каждом узле которой помещен атом массы m.

На рис. III.12 изображена простая кубическая решетка и прямоугольная координатная система x, y, z с началом в атоме O. Шесть ближайших ●-атомов расположены на расстоянии a от центрального O-атома (*первая координационная группа*). Следующие 12 атомов, расположенные наиболее близко к O (*вторая координационная группа*), находятся от O на расстоянии $a\sqrt{2}$. Из них атомы *1*, *2*, *3*, *4* лежат в плоскости xy, атомы *5*, *6*, *7*, *8* — в плоскости yz и атомы *9*, *10*, *11*, *12* — в плоскости xz[2]).

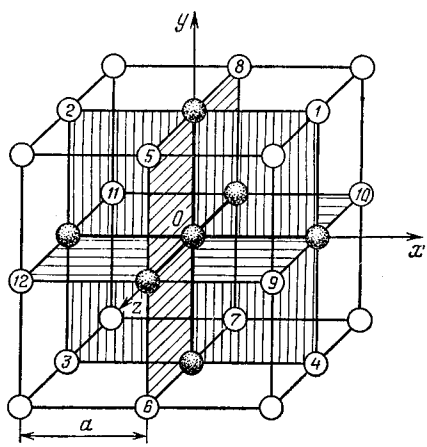


Рис. III.12.

Для упрощения записи обозначим вектор прямой решетки

$$\boldsymbol{a}_n=n_1\boldsymbol{a}_1+n_2\boldsymbol{a}_2+n_3\boldsymbol{a}_3\equiv\boldsymbol{n}$$
$$(a_1=a_2=a_3=a). \quad (7.1)$$

Из (5.2) следует, что сила, действующая на атом $(\boldsymbol{n},\ k)$ в направлении α, при смещении атома $(\boldsymbol{n}',\ k')$ в направлении β

[1]) При суммировании в (6.12) по половине бриллюэновской зоны следует опустить множитель 1/2.

[2]) Учет взаимодействия со второй координационной группой необходим для обеспечения стабильности решетки.

на величину $u^{k'}_{n'\beta}$, в линейном приближении равна

$$-\frac{\partial \Phi}{\partial u^{k}_{n\alpha}} = -\Phi_{\alpha\beta}\begin{pmatrix} k & k' \\ \boldsymbol{n} & \boldsymbol{n}' \end{pmatrix} u^{k'}_{n'\beta}. \tag{7.2}$$

В случае простой решетки индексы k и k' отсутствуют и сила, действующая на атом в начале координат ($\boldsymbol{n}=0$) в направлении оси α, при смещении атома $\boldsymbol{n}'$ ($\equiv \boldsymbol{n}$) в направлении β на величину $u_{n\beta}$, равна

$$-\frac{\partial \Phi}{\partial u_{0\alpha}} = -\Phi^{0n}_{\alpha\beta} u_{n\beta}. \tag{7.3}$$

Сила (7.3) может быть выражена через константу квазиупругой связи α_n, зависящую, вообще говоря, от расстояния между атомами $n=|\boldsymbol{n}|$. Имеем

$$-\Phi^{0n}_{\alpha\beta} u_{n\beta} = \alpha_n u_{n\beta} \frac{n_\beta}{n} \frac{n_\alpha}{n}, \tag{7.4}$$

где n_β и n_α — проекции расстояния n на оси β и α. Отсюда

$$\Phi^{0n}_{\alpha\beta} = -\alpha_n \frac{n_\alpha n_\beta}{n^2}. \tag{7.5}$$

При определении $D_{\alpha\beta}(\boldsymbol{q})$ по формуле (5.10а) необходимо суммировать по всем $\boldsymbol{n}'$, в том числе и по $\boldsymbol{n}'=0$; в то же время формулы (7.4) и (7.5) применимы только в случае $\boldsymbol{n}\neq 0$ (мы заменили $\boldsymbol{n}'$ на $\boldsymbol{n}$).

Для того чтобы определить $\Phi^{00}_{\alpha\beta}$, воспользуемся соотношением (5.3б), откуда

$$\Phi^{00}_{\alpha\beta} = -\sum_{(\boldsymbol{n}\neq 0)} \Phi^{0n}_{\alpha\beta}. \tag{7.6}$$

Из (5.10а), (7.5) и (7.6) следует для динамической матрицы

$$D_{\alpha\beta}(\boldsymbol{q}) = \frac{1}{m} \sum_{(\boldsymbol{n}\neq 0)} \alpha_n \frac{n_\alpha n_\beta}{n^2} (1 - e^{i\boldsymbol{q}\boldsymbol{n}}), \tag{7.7}$$

где m — масса атома.

Вычислим $D_{xx}(\boldsymbol{q})$. Легко видеть, что произведение $n_x n_x = n_x^2$ отлично от нуля только для двух ●-атомов, лежащих на оси x (рис. III.12), поэтому для первой координационной группы

$$D^{(1)}_{xx}(\boldsymbol{q}) = \frac{1}{m} \sum_{(\boldsymbol{n}\neq 0)} \alpha_1 \frac{n_x^2}{n^2} (1 - e^{i\boldsymbol{q}\boldsymbol{n}}) =$$

$$= \frac{\alpha_1}{m} [1 - e^{iq_x a} + 1 - e^{-iq_x a}] = \frac{2\alpha_1}{m} (1 - \cos q_x a) = \frac{4\alpha_1}{m} \sin^2 \frac{q_x a}{2}. \tag{7.8}$$

Четыре атома второй координационной группы (*5*, *6*, *7*, *8*; рис. III.12) вклада в $D^{(1)}_{xx}$ не дают, так как для них $n_x = 0$. Учиты-

вая вклад восьми остающихся атомов второй координационной группы, получим

$$D_{xx}^{(2)}(\boldsymbol{q}) = \frac{1}{m}\sum_{(\boldsymbol{n}\neq 0)} \alpha_2 \frac{n_x^2}{n^2}(1-e^{i\boldsymbol{q}\boldsymbol{n}}) =$$

$$= \frac{\alpha_2}{m}\left(\frac{1}{\sqrt{2}}\right)^2 [8-2\cos(q_x a+q_y a)-2\cos(q_x a-q_y a)-$$

$$-2\cos(q_x a+q_z a)-2\cos(q_x a-q_z a) =$$

$$= \frac{2\alpha_2}{m}[2-\cos q_x a\cos q_y a-\cos q_x a\cos q_z a]. \quad (7.9)$$

При вычислении этой суммы, например, атомы *1* и *3* (рис. III.12) дают

$$(1-e^{i(q_x a+q_y a)})+(1-e^{-i(q_x a+q_y a)}) = 2-2\cos(q_x a+q_y a),$$

аналогично другие пары атомов.

Из (7.8) и (7.9) следует, что

$$D_{xx}(\boldsymbol{q}) = D_{xx}^{(1)}(q)+D_{xx}^{(2)}(q) =$$

$$= \frac{4\alpha_1}{m}\sin^2\frac{q_x a}{2}+\frac{2\alpha_2}{m}[2-\cos q_x a(\cos q_y a+\cos q_z a)]. \quad (7.10)$$

Остальные диагональные члены динамической матрицы $D_{yy}(\boldsymbol{q})$ и $D_{zz}(\boldsymbol{q})$ получаются из (7.10) посредством циклической перестановки индексов x, y и z.

Недиагональные члены динамической матрицы, например,

$$D_{xy}(\boldsymbol{q}) = \frac{1}{m}\sum_{(\boldsymbol{n}\neq 0)} \alpha_n \frac{n_x n_y}{n^2}(1-e^{i\boldsymbol{q}\boldsymbol{n}}). \quad (7.11)$$

Атомы первой координационной группы вклада в $D_{xy}(\boldsymbol{q})$ не вносят, так как для них либо n_x, либо n_y равны нулю. Из второй координационной группы вклад вносят только четыре атома, лежащие в плоскости xy:

$$D_{xy}(\boldsymbol{q}) = \frac{\alpha_2}{m}\left(\frac{1}{\sqrt{2}}\right)^2 [(1-e^{i(q_x a+q_y a)})-$$

$$-(1-e^{i(q_y a-q_x a)})+(1-e^{-i(q_x a+q_y a)})-(1-e^{-i(q_y a-q_x a)}) =$$

$$= \frac{\alpha_2}{2m}[2\cos(q_y a-q_x a)-2\cos(q_x a+q_y a)] =$$

$$= \frac{2\alpha_2}{m}\sin q_x a\sin q_y a, \quad (7.12)$$

$$D_{xy}(\boldsymbol{q}) = \frac{2\alpha_2}{m}\sin q_x a\sin q_y a. \quad (7.13)$$

Остальные недиагональные члены D_{xz}, D_{yz} и т. д. могут быть получены циклической перестановкой.

Из (7.10) и (7.13) мы видим, что элементы динамической матрицы вещественны и симметричны, т. е. $D_{\alpha\beta}=D_{\beta\alpha}$, как это и должно быть для простых решеток (см. (5.11в)). Характеристическое уравнение (5.11) имеет вид

$$|D_{\alpha\beta}(\boldsymbol{q})-\omega^2\delta_{\alpha\beta}|=0, \quad (7.14)$$

где $D_{\alpha\beta}(\boldsymbol{q})$ определяются выражениями (7.10) и (7.13).

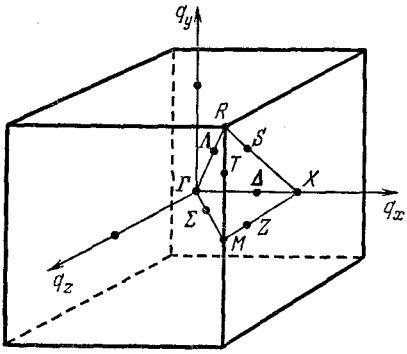


Рис. III. 13.

Уравнение (7.14) третьей степени относительно ω^2, поэтому его решение в общем случае затруднительно; мы его решим для избранных симметричных точек (линий) зоны Бриллюэна.

Для простой кубической решетки зона Бриллюэна имеет форму куба, изображенную на рис. III.13.

Определим вид спектра (дисперсию) и направление векторов поляризации e_j ($j=1, 2, 3$) для точек Δ, т. е. оси симметрии [100]. В этом случае $q_x=q$, $q_y=q_z=0$, и из (7.10) и (7.13) следует, что

$$D_{xx}=\frac{4(\alpha_1+2\alpha_2)}{m}\sin^2\frac{qa}{2}, \quad D_{yy}=D_{zz}=\frac{4\alpha_2}{m}\sin^2\frac{qa}{2}. \quad (7.15)$$

Все недиагональные члены D_{xy}, D_{xz} и т. д. равны нулю поэтому характеристическое уравнение имеет вид

$$(D_{xx}-\omega^2)(D_{yy}-\omega^2)^2=0, \quad (7.16)$$

откуда три ($j=1, 2, 3$) его корня:

$$\omega_1^2=D_{xx}=\frac{4(\alpha_1+2\alpha_2)}{m}\sin^2\frac{qa}{2}, \quad \omega_2^2=\omega_3^2=\frac{4\alpha_2}{m}\sin^2\frac{qa}{2}. \quad (7.17)$$

Уравнения для поляризационных векторов $\boldsymbol{e}_j$ имеют вид (см. (6.1)):

$$\left.\begin{aligned} D_{xx}e_{1x}&=\omega_1^2e_{1x},\\ D_{yy}e_{1y}&=\omega_1^2e_{1y},\\ D_{zz}e_{1z}&=\omega_1^2e_{1z},\end{aligned}\right\} \quad (7.18)$$

$$\left.\begin{aligned} D_{xx}e_{2x}&=\omega_2^2e_{2x},\\ D_{yy}e_{2y}&=\omega_2^2e_{2y},\\ D_{zz}e_{2z}&=\omega_2^2e_{2z}.\end{aligned}\right\} \quad (7.19)$$

Уравнения для вектора $\boldsymbol{e}_3$ совпадают с (7.19). Решения однородных уравнений (7.18) и (7.19), с учетом нормировки $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ и $\boldsymbol{e}_3$ (см. (6.2) и (6.3)), имеют вид: $\boldsymbol{e}_1(1, 0, 0)$, $\boldsymbol{e}_2(0, 1, 0)$ и $\boldsymbol{e}_3(0, 0, 1)$; из-за вырождения частоты $\omega_2=\omega_3$ выбор составляющих у $\boldsymbol{e}_2$ и $\boldsymbol{e}_3$ неоднозначен, можно, например, положить

$\boldsymbol{e}_2(0, -1/\sqrt{2}, 1/\sqrt{2})$ и $\boldsymbol{e}_3(0, 1/\sqrt{2}, 1/\sqrt{2})$, что, однако, не приводит к новым физически интересным результатам.

Таким образом, волна распространяющаяся вдоль оси x с волновым вектором $q = q_x$, характеризуется продольным вектором поляризации $\boldsymbol{e}_L(1, 0, 0)$ и поперечными векторами поляризации $\boldsymbol{e}_T^{(1)}(0, 1, 0)$ и $\boldsymbol{e}_T^{(2)}(0, 0, 1)$.

Секулярное уравнение (7.14) может быть просто решено и для других симметричных точек (линий) Λ, Σ и Z (рис. III.13). Определив для этих точек соответствующие значения ω_j^2, можно из уравнений (6.1) определить поляризационные векторы. Так как процедура эта подобна проделанной нами выше для точки Δ, мы не воспроизводим вычислений для остальных точек, ограничиваясь приведением результатов в табл. III.1.

Рассматривая табл. III.1, видим, что поляризационные векторы $\boldsymbol{e}_j$ нормированы на единицу в случаях а) и г). В двух

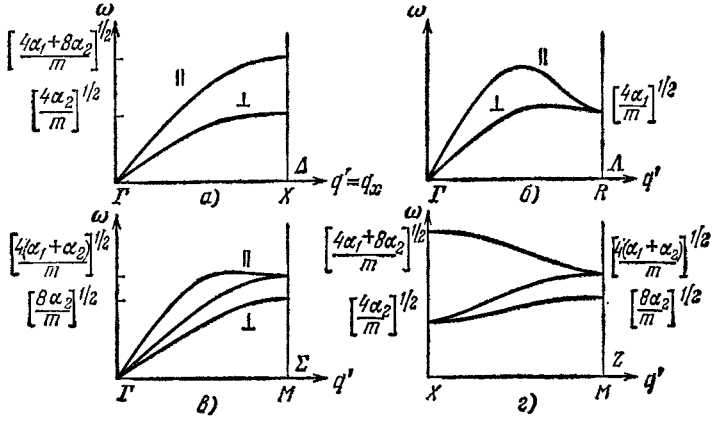


Рис. III.14.

других случаях, мы для наглядности предпочли положить равными ± 1, все не равные нулю составляющие у $\boldsymbol{e}_j$. В случае б) векторы $\boldsymbol{e}_T^{(1)}$ и $\boldsymbol{e}_T^{(2)}$ не ортогональны друг к другу (они образуют угол в 80°, в чем можно убедиться, составив их скалярное произведение). Это не противоречит общей теореме, доказанной в Приложении 3 п. 5, так как собственные векторы $e_T^{(1)}$ и $e_T^{(2)}$ относятся к одному собственному значению $\omega_2^2 = \omega_3^2$. Можно, конечно, и в этом случае посредством линейного преобразования ввести взаимно ортогональные собственные векторы. В случае г) ветви колебаний не разделяются на продольные и поперечные (только $\boldsymbol{e}_2(0, 1, 0)$ — поперечное колебание).

На рис. III.14 схематически изображены зависимости частот ω_j от q' в случаях а) — г). В случаях *а*), *б*), *в*) линии симметрии

Таблица III.1

а) Линия Δ: $q_x = q$, $q_y = q_z = 0$.

$$\|D_{\alpha\beta}\| = \begin{Vmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & B \end{Vmatrix}, \qquad A = \frac{4\alpha_1 + 8\alpha_2}{m} \sin^2 \frac{qa}{2}, \quad B = \frac{4\alpha_2}{m} \sin^2 \frac{qa}{2},$$

$\omega_1^2 = A$, $\quad \boldsymbol{e}_L\,(1, 0, 0)$

$\omega_2^2 = \omega_3^2 = B$, $\quad \boldsymbol{e}_T^{(1)}\,(0, 1, 0)$, $\boldsymbol{e}_T^{(2)}\,(0, 0, 1)$;

б) Линия Λ: $q_x = q_y = q_z = q' = \dfrac{q}{\sqrt{3}}$.

$$\|D_{\alpha\beta}\| = \begin{Vmatrix} A & B & B \\ B & A & B \\ B & B & A \end{Vmatrix}, \quad A = \frac{4\alpha_1}{m} \sin^2 \frac{q'a}{2} + \frac{4\alpha_2}{m} \sin^2 q'\alpha, \quad B = \frac{2\alpha_2}{m} \sin^2 q'a,$$

$\omega_1^2 = A + 2B$, $\quad \boldsymbol{e}_L\,(1, 1, 1)$,

$\omega_2^2 = \omega_3^2 = A - B$, $\quad \boldsymbol{e}_T^{(1)}\,(1, -1, 0)$, $\boldsymbol{e}_T^{(2)}\,(1, 0, -1)$.

в) Линия Σ: $q_x = q'$, $q_y = 0$, $q_z = q' = \dfrac{q}{\sqrt{2}}$.

$$\|D_{\alpha\beta}\| = \begin{Vmatrix} A & 0 & C \\ 0 & B & 0 \\ C & 0 & A \end{Vmatrix}, \quad A = \frac{4(\alpha_1 + \alpha_2)}{m} \sin^2 \frac{q'a}{2} + \frac{2\alpha_2}{m} \sin^2 q'\boldsymbol{a},$$

$$B = \frac{8\alpha_2}{m} \sin^2 \frac{q'a}{2}, \quad C = \frac{2\alpha_2}{m} \sin^2 q'a.$$

$\omega_1^2 = A + C$, $\quad \boldsymbol{e}_L\,(1, 0, 1)$,

$\omega_2^2 = B$, $\quad \boldsymbol{e}_T^{(1)}\,(0, 1, 0)$,

$\omega_3^2 = A - C$, $\quad \boldsymbol{e}_T^{(2)}\,(1, 0, -1)$.

г) Линия Z: $q_x = \dfrac{\pi}{a}$, $q_y = 0$, $q_z = q'' = \sqrt{q^2 - \left(\dfrac{\pi}{a}\right)^2}$.

$$\|D_{\alpha\beta}\| = \begin{Vmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{Vmatrix}, \quad A = \frac{4\alpha_1}{m} + \frac{2\alpha_2}{m} \{3 + \cos q'a\},$$

$$B = \frac{2\alpha^2}{m} \{3 - \cos q'\boldsymbol{a}\}, \quad C = \frac{4\alpha_1}{m} \sin^2 \frac{q'a}{2} + \frac{4\alpha_2}{m}.$$

$\omega_1^2 = A$, $\quad \boldsymbol{e}_1\,(1, 0, 0)$,

$\omega_2^2 = B$, $\quad \boldsymbol{e}_2\,(0, 1, 0) \equiv \boldsymbol{e}_T$,

$\omega_3^2 = C$, $\quad \boldsymbol{e}_3\,(0, 0, 1)$.

выходят из точки Г — центра бриллюэновской зоны, которому соответствует $\boldsymbol{q}=0$, поэтому акустические ветви в этом случае, как и должно быть, начинаются с $\omega_j=0$. Как следует из табл. III.1, в этих случаях, для малых значений волнового вектора $\boldsymbol{q}$, $\omega=v_0 q$, где v_0 — скорость звука (скорость распространения длинных волн: $qa=2\pi a/\lambda \ll 1$, т. е. длина волны $\lambda \gg a=$ постоянной решетки).

Из табл. III.1 легко определить скорости продольной и поперечной волн звука, во всех трех случаях (*а*, *б*, *в*). Например, в случае *а*)

$$v_{0L}=a\sqrt{\frac{\alpha_1+2\alpha_2}{m}}, \quad v_{0T}=a\sqrt{\frac{\alpha_2}{m}},$$

откуда следует, что $v_{0L}>v_{0T}$; это имеет место во всех трех случаях (в случаях *б*) и *в*) надо положить $q'=q/\sqrt{3}$ и $q'=q/\sqrt{2}$).

Линия симметрии Z не проходит через центр бриллюэновской зоны (рис. III.13), поэтому дисперсионные кривые $\omega_j=\omega_j(q)$ не начинаются с нуля. Заметим, что $q_x=q=\pi/a$ в случае *а*) и $q_z=q'=0$ в случае *г*) соответствует одна и та же точка в бриллюэновской зоне X, поэтому предельные значения частот в обоих случаях совпадают.

§ 8. Применение теории групп к исследованию нормальных колебаний кристаллической решетки

1. В (гл. II, § 7) мы исследовали в общем виде вопрос о применении теории групп к квантовомеханической системе, гамильтониан которой обладает некоторой симметрией. В первом пункте настоящего параграфа мы рассмотрим аналогичный вопрос в применении к нормальным колебаниям кристаллической решетки; затем мы применим полученные результаты к простой кубической решетке, исследованной в предыдущем параграфе.

Будем исходить из понятия группы волнового вектора $G_{\boldsymbol{q}}$, введенной в гл. II, § 8, п. 3. Группа волнового вектора $G_{\boldsymbol{q}}$ есть подгруппа пространственной группы кристалла G, имеющая элементы $g=\{R\,|\,\boldsymbol{\alpha}(R)+\boldsymbol{a}_n\}$ такие, что $R\boldsymbol{q}=\boldsymbol{q}$ или $R\boldsymbol{q}=\boldsymbol{q}+\boldsymbol{b}_i$, где $\boldsymbol{b}_i$ — вектор обратной решетки.

Подействуем оператором $\hat{P}_g$ (II.6.1), где $g\in G_{\boldsymbol{q}}$, на обе части уравнения (6.1). Применение оператора $\hat{P}_g$ к динамической матрице $D^{kk'}_{\alpha\beta}(\boldsymbol{q})$ (5.10а) оставляет ее без изменения. В самом деле, поскольку элементы g являются элементами симметрии кристалла ($g\in G_{\boldsymbol{q}}\in G$), а силовые коэффициенты $\Phi_{\alpha\beta}\begin{pmatrix}k\,k'\\ \boldsymbol{n}\boldsymbol{n}'\end{pmatrix}$ берутся для равновесных положений атомов в решетке, преобразование $\hat{P}_g$ может изменить только нумерацию узлов решетки $\boldsymbol{n}$, $\boldsymbol{n}'$, по которым производится суммирование при определении $D^{kk'}_{\alpha\beta}(\boldsymbol{q})$ ($\boldsymbol{n}$ можно

всегда положить равным нулю); что же касается волнового вектора $\boldsymbol{q}$, то он не меняется по определению группы волнового вектора G_q. В результате мы из (6.1) получим

$$\sum_{k'\beta} D^{kk'}_{\alpha\beta}(q)\,\hat{P}_g e_{jk'\beta}(q) = \omega_j^2(q)\,\hat{P}_g e_{jk\alpha}(q). \tag{8.1}$$

Это уравнение в некотором отношении аналогично уравнению Шредингера (II.7.13): роль, гамильтониана $\hat{\mathcal{H}}(\boldsymbol{x})$ играет динамическая матрица $D^{kk'}_{\alpha\beta}(\boldsymbol{q})$, роль собственных функций $\psi(\boldsymbol{x})$ — собственные векторы $e_{jk\alpha}(\boldsymbol{q})$ и роль собственных значений энергии $\mathscr{E}$ — нормальные частоты $\omega_j^2(\boldsymbol{q})$.

Мы видим, что собственные векторы $e_{jk\alpha}(q)$ играют роль базисных векторов неприводимого представления группы G_q. Если собственное значение $\omega_j^2(\boldsymbol{q})$ уравнения (6.1) вырождено, т. е. имеется несколько линейно-независимых собственных векторов $e_{jk\alpha}(\boldsymbol{q})$, соответствующих частоте $\omega_j(\boldsymbol{q})$, то собственный вектор $\hat{P}_g e_{jk\alpha}(\boldsymbol{q})$ может быть представлен как линейная комбинация собственных векторов соответствующих частоте $\omega_j(\boldsymbol{q})$. Матрицы этого линейного представления, аналогично (II.7.15), осуществляют неприводимое представление группы волнового вектора G_q. Размерность неприводимого представления равна кратности вырождения частоты $\omega_j(\boldsymbol{q})$.

Для того чтобы применить теорию групп к классификации колебаний кристаллов в разных точках зоны Бриллюэна, надо аналогично тому, как это было сделано для молекул (гл. III, § 8, п. 3), определить полное (приводимое) представление, соответствующее колебательным степеням свободы для определенного $\boldsymbol{q}$. Для сложного кристалла полное число колебательных степеней свободы $u^k_{n\alpha}$ равно $3Ns-6\approx 3Ns$ (N — число кристаллических ячеек в основной области, s — число атомов в ячейке), т. е. очень велико.

Однако нас интересует классификация колебаний в данной точке бриллюэновской зоны, т. е. для определенного значения волнового вектора $\boldsymbol{q}$, который (для одной ветви колебаний) сам принимает N квазидискретных значений (5.24). Поэтому число колебательных степеней свободы при данном значении $\boldsymbol{q}$ равно $3s$, т. е. равно числу степеней свободы для атомов элементарной ячейки.

Мы должны от $3sN$ смещений атомов $u^k_{n\alpha}$ перейти к их фурье-образам $u^k_{n\alpha}(\boldsymbol{q})$ (как это сделано ниже) и определить, по каким представлениям пространственной группы волнового вектора G_q они преобразуются.

Будем исходить из выражения (6.4), которое мы запишем в виде

$$u^k_{n\alpha} = \sum_{q} u_{k\alpha}(\boldsymbol{q})\, e^{iqa_n}, \tag{8.2}$$

где

$$u_{k\alpha}(\boldsymbol{q}) = \frac{1}{\sqrt{Nm_k}} \sum_j e_{jk\alpha}(\boldsymbol{q})\, a_j(\boldsymbol{q},\ t). \tag{8.2a}$$

Здесь $\boldsymbol{e}_{jk}$ — поляризационные векторы, a_j — комплексные нормальные координаты.

Умножая обе части (8.2) на $\exp(-i\boldsymbol{q}'\boldsymbol{a}_n)$, суммируя по $\boldsymbol{n}$ и используя (П.6.4), получим

$$u_{k\alpha}(\boldsymbol{q}) = \frac{1}{N} \sum_{\boldsymbol{n}} e^{-iq a_{\boldsymbol{n}}} u^k_{n\alpha} \tag{8.3}$$

(где мы в последнем выражении обозначили $\boldsymbol{q}'$ через $\boldsymbol{q}$).

Применим к $3s$ (при заданном значении $\boldsymbol{q}$) величинам $u_{k\alpha}(\boldsymbol{q})$ оператор $\hat{P}_{g^{-1}}$ (II.6.1), где

$$g = \{R \mid \boldsymbol{a}_m + \boldsymbol{\alpha}\} \tag{8.4}$$

— элемент группы волнового вектора G_q (очевидно, g^{-1} — тоже элемент группы волнового вектора G_q).

Из (8.3) следует

$$\hat{P}_{g^{-1}} u_{k\alpha}(\boldsymbol{q}) = \frac{1}{N} \sum_{\boldsymbol{n}} e^{-iq a_{\boldsymbol{n}}} \hat{P}_{g^{-1}} u^k_{n\alpha}. \tag{8.5}$$

Оператор $\hat{P}_{g^{-1}}$ действует в правой части (8.3) только на смещения $u^k_{n\alpha}$, так как множители $\exp(-i\boldsymbol{q}\boldsymbol{a}_n)$ являются коэффициентами разложения величины $u_{k\alpha}(\boldsymbol{q})$ по смещениям $u^k_{n\alpha}$ (аналогично в гл. II, § 8, п. 3) оператор $\hat{P}_R$ преобразует смещения $u_{i\alpha}$ атомов в молекуле, не действуя на нумерацию атомов).

Так как $u^k_{n\alpha}$ — составляющая вектора смещения, то аналогично (II.8.13)

$$\hat{P}_{g^{-1}} u^k_{n\alpha} = \sum_{n'k'\alpha'} D^u(g^{-1})_{n'k'\alpha',\ nk\alpha} u^{k'}_{n'\alpha'}. \tag{8.6}$$

Матрица D^u размерности $3sN$ осуществляет представление смещений всех атомов кристалла. Из (II.8.13) следует, что

$$D^u(g^{-1})_{n'k'\alpha',\ nk\alpha} = A(g^{-1})_{\alpha'\alpha} \delta_{n',\ g(n)} \delta_{k',\ g(k)}. \tag{8.7}$$

Здесь $A(g^{-1})$ — матрица преобразований компонент полярного вектора $u^k_{n\alpha}$ при преобразовании g^{-1} координатной системы; символы Кронекера $\delta_{n',\ g(n)}$, $\delta_{k',\ g(k)}$ учитывают, что в результате операции g атом $(\boldsymbol{n},\ k)$ переходит на место атома того же сорта $(\boldsymbol{n}',\ k')$.

Подставляя (8.7) и (8.6) в (8.5) и преобразуя тождественно экспоненту, получим

$$\hat{P}_{g^{-1}} u_{k\alpha}(\boldsymbol{q}) = \frac{1}{N} \sum_{n,\ n',\ k',\ \alpha'} e^{-iq a'_n} u^{k'}_{n'\alpha'} A(g^{-1})_{\alpha'\alpha} e^{iq(a'_n - a_n)} \delta_{n',\ g(n)} \delta_{k',\ g(k)}. \tag{8.8}$$

Так как операции g преобразует $\boldsymbol{a}_n$ в $\boldsymbol{a}'_n$, то

$$i\boldsymbol{q}(\boldsymbol{a}'_n-\boldsymbol{a}_n)=i\boldsymbol{q}(g\boldsymbol{a}_n-\boldsymbol{a}_n)=i\boldsymbol{q}(R\boldsymbol{a}_n+\boldsymbol{a}_m+\boldsymbol{\alpha}-\boldsymbol{a}_n)= \\ =i[(R^{-1}\boldsymbol{q}-\boldsymbol{q})\boldsymbol{a}_n+\boldsymbol{q}(\boldsymbol{a}_m+\boldsymbol{\alpha})]. \quad (8.9)$$

Если g^{-1}—элемент группы волнового вектора G_q, то

$$R^{-1}\boldsymbol{q}=\boldsymbol{q} \text{ или } R^{-1}\boldsymbol{q}=\boldsymbol{q}\pm\boldsymbol{b}_i \quad (8.10)$$

($\boldsymbol{b}_i$ — вектор обратной решетки). $\exp[i(\boldsymbol{R}^{-1}\boldsymbol{q}-\boldsymbol{q})\boldsymbol{a}_n]=1$ в обоих случаях, поэтому

$$\hat{P}_{g^{-1}}u_{k\alpha}(\boldsymbol{q})= \\ =\frac{1}{N}\sum_{n,\,n',\,k'\alpha'}e^{-iqa'_n}u^{k'}_{n'\alpha'}A(g^{-1})_{\alpha'\alpha}e^{iq(a_m+\alpha)}\delta_{n',\,g(n)}\delta_{k',\,g(k)}. \quad (8.11)$$

Если просуммировать в правой части по $\boldsymbol{n}'$ и воспользоваться (8.3), то получим

$$\hat{P}_{g^{-1}}u_{k\alpha}(\boldsymbol{q})=\sum_{k',\,\alpha'}A(g^{-1})_{\alpha'\alpha}e^{iq(a_m+\alpha)}\delta_{k',\,g(k)}u_{k'\alpha'}(\boldsymbol{q}) \quad (8.12)$$

(суммирование по $\boldsymbol{n}$ выполняется автоматически из-за наличия множителя $\delta_{n',\,g(n)}$).

Покажем, что для симморфных групп волнового вектора G_q, а также для внутренних точек зоны Бриллюэна любых групп G_q матрицы $A(g^{-1})\exp[i\boldsymbol{q}(\boldsymbol{\alpha}+\boldsymbol{a}_m)]\delta_{k',\,g(k)}$ реализуют на базисных векторах $u_{k\alpha}(\boldsymbol{q})$ представления групп волнового вектора G_q размерности $3s$. В самом деле, так как указанные выше матрицы содержат множитель $\exp[i\boldsymbol{q}(\boldsymbol{\alpha}+\boldsymbol{a}_m)]$, то произведение матриц $g_1=\{R_1|\boldsymbol{\alpha}_1+\boldsymbol{a}_1\}$ и $g_2=\{R_2|\boldsymbol{\alpha}_2+\boldsymbol{a}_2\}$ для элементов группы G_q будет содержать множитель $\exp[i\boldsymbol{q}(\boldsymbol{\alpha}_1+\boldsymbol{a}_1+\boldsymbol{\alpha}_2+\boldsymbol{a}_2)]$.

В то время, как мы показали в (II. § 9 п.4), что неприводимое представление $\Gamma(g_2g_1)$, в том случае содержит множитель: $\exp[i\boldsymbol{q}(R_2\boldsymbol{\alpha}_1+R_2\boldsymbol{a}_1+\boldsymbol{\alpha}_2+\boldsymbol{a}_2)]$ (II.9.31); обе экспоненты совпадают, если $\exp[i\boldsymbol{q}(\boldsymbol{\alpha}_1+\boldsymbol{a}_1)]=\exp[i\boldsymbol{q}(R_2\boldsymbol{\alpha}+R_2\boldsymbol{a}_1)]$, т. е. $\exp[i(R_2^{-1}\boldsymbol{q}-\boldsymbol{q})+(\boldsymbol{\alpha}_1+\boldsymbol{a}_1)]=1$, что, учитывая (8.10), выполняется либо для симморфных групп ($\boldsymbol{\alpha}_1=0$), либо для внутренних точек зоны Бриллюэна ($R_2^{-1}\boldsymbol{q}=\boldsymbol{q}$).

Раскладывая это представление размерности $3s$ по неприводимым представлениям группы G_q, мы можем классифицировать колебания кристалла в точке $\boldsymbol{q}$ его бриллюэновской зоны. Отметим близкую аналогию выражения (8.12) с (II.8.13), полученным при анализе колебаний многоатомных молекул.

Так как для разложения представления (8.7) по неприводимым представлениям группы волнового вектора G_q достаточно знать соответствующие характеры, определим характеры представления (8.7). Воспользовавшись (II.8.15), получим при собст-

венном вращении на угол φ для характера представления D^u (8.7):

$$\chi^u(C(\varphi)\,|\,\alpha) = (1+2\cos\varphi)\exp(i\boldsymbol{q}\boldsymbol{\alpha})\,n_C, \tag{8.13}$$

где $n_C = \sum_k \delta_{k,\,C(k)}$ — число атомов ячейки, остающихся на месте, при преобразовании $C(\varphi) \equiv C$.

Аналогично для зеркально-поворотного преобразования $S(\varphi) \equiv S$ (II.8.18):

$$\chi^u(S(\varphi)\,|\,\alpha) = (-1+2\cos\varphi)\exp(i\boldsymbol{q}\boldsymbol{\alpha})\,n_S, \tag{8.14}$$

где n_S — число атомов ячейки, остающихся на месте при преобразовании $S(\varphi)$.

2. Применим теперь полученные нами результаты к исследованию колебаний в простом кубическом кристалле, рассмотренном в предыдущем параграфе.

Определим неприводимые представления группы волнового вектора G_q в центре бриллюэновской зоны ($\boldsymbol{q}=0$), т. е. в точке Г (рис. III.13). Точка Г обладает симметрией куба $\boldsymbol{O}_h = \boldsymbol{O}\times\boldsymbol{C}_i$; здесь $\boldsymbol{O}$ — группа симметрии осей куба, а $\boldsymbol{C}_i = \{E, J\}$, где J — инверсия (гл. II, § 3). Табл. III.2[1]) характеров группы $\boldsymbol{O}_h$ может быть получена по схеме табл. II.6, если воспользоваться табл. II.7.

Для определения характеров представления D^u_Γ (8.7) в точке Г воспользуемся выражением (8.13), в котором следует положить $\boldsymbol{\alpha}=0$ и $n_C=1$. Кроме того, учтем, что для элементов, содержащих инверсию J,

$$\chi^u[JC(\varphi)\,|\,\alpha] = -(1+2\cos\varphi)\exp(i\boldsymbol{q}\boldsymbol{\alpha})\,n_{CJ}, \tag{8.15}$$

как это следует из предыдущего вывода и из (II.8.14), если заменить в правой части все u_α на $-u_\alpha$.

В результате мы получим

$$\begin{gathered}\chi^u(E)=3,\quad \chi^u(C_3)=0,\quad \chi^u(C_4^2)=-1,\quad \chi^u(C_2)=-1,\\ \chi^u(C_4)=1,\quad \chi^4(J)=-3,\quad \chi^u(JC_3)=0,\quad \chi^u(JC_4^2)=1,\\ \chi^u(JC_2)=1,\quad \chi^u(JC_4)=-1.\end{gathered} \tag{8.16}$$

Мы видим, что характеры представления D^u_Γ совпадают с характерами неприводимого представления Γ_{15} (табл. III.2), но

[1]) Обозначения неприводимых представлений в табл. III.2 заимствованы из статьи Баукарта Л. П., Смолуховского Р., Вигнера Е. Теория зон Бриллюэна и свойства симметрии волновых функций в кристаллах./ В кн. Нокс Р., Голд А. Симметрия в твердом теле.— М.: Наука, 1970, с. 187. Эти обозначения неприводимых представлений оправданы, по мнению Г. Джонса, «... так как они настолько установились в литературе, что их использование представляется неизбежным».

тогда

$$D^{u}_{\Gamma} = \Gamma_{15}, \tag{8.17}$$

т. е. группе волнового вектора в центре бриллюэновской зоны соответствует трехкратно вырожденное неприводимое представление Γ_{15}.

Для определения группы волнового вектора в точке Δ (см. рис. III.13) рассмотрим, под действием каких элементов группы вектор $\boldsymbol{q}_{\Delta}$ остается без изменения. Легко видеть, что это будут элементы группы $\boldsymbol{C}_{4v}$: E, C_4^2 (ось q_x), $2C_4$ (q_x), $2JC_4^2$ (оси q_y и q_z), $2JC_2$; заметим, что операции JC_4^2 и JC_2 — отражения в плоскостях, перпендикулярных к осям C_4 и C_2, проходящих через ось q_x. Точка T зоны Бриллюэна удовлетворяет той же группе точечной симметрии $\boldsymbol{C}_{4v}$ (здесь надо учесть, что на четырех вертикальных ребрах куба имеются четыре эквивалентные точки T). Составим теперь по общему рецепту табл. III.3 характеров группы $G_{q_{\Delta}}$ и G_{q_T}. Аналогично могут быть составлены таблицы характеров групп волновых векторов, направленных к точкам Λ и Σ (точка S имеет ту же симметрию, что и точка Σ) (табл. III.4).

Таблица III.2

O_h	E	$8C_3$	$3C_4^2$	$6C_2$	$6C_4$	J	$8JC_3$	$3JC_4^2$	$6JC_2$	$6JC_4$
Γ_1	1	1	1	1	1	1	1	1	1	1
Γ_2	1	1	1	−1	−1	1	1	1	−1	−1
Γ_{12}	2	−1	2	0	0	2	−1	2	0	0
Γ'_{15}	3	0	−1	−1	1	3	0	−1	−1	1
Γ'_{25}	3	0	−1	1	−1	3	0	−1	1	−1
Γ'_1	1	1	1	1	1	−1	−1	−1	−1	−1
Γ'_2	1	1	1	−1	−1	−1	−1	−1	1	1
Γ'_{12}	2	−1	2	0	0	−2	1	−2	0	0
Γ_{15}	3	0	−1	−1	1	−3	0	1	1	−1
Γ_{25}	3	0	−1	1	−1	−3	0	1	−1	1

Таблица III.3

	Δ, T	E	C_4^2	$2JC_2$	$2JC_4^2$	$2C_4$
	Δ_1	1	1	1	1	1
	Δ_1'	1	1	−1	−1	1
	Δ_2	1	1	−1	1	−1
	Δ_2'	1	1	1	−1	−1
	Δ_5	2	−2	0	0	0
χ^u	Γ_{15}	3	−1	1	1	1

Таблица III.4

Λ	E	$2C_3$	$3JC_2$
Λ_1	1	1	1
Λ_2	1	1	−1
Λ_3	2	−1	0
Γ_{15}	3	0	1

а)

Σ, S	E	C_2	JC_4^2	JC_2
Σ_1	1	1	1	1
Σ_2	1	1	−1	−1
Σ_3	1	−1	−1	1
Σ_4	1	−1	1	−1
Γ_{15}	3	−1	1	1

б)

Неприводимые представления группы волнового вектора $\boldsymbol{q}_\Delta$ могут быть определены двумя способами. Во-первых, мы можем определить характеры представления D^u_Δ (8.7) для группы C_{4v}:

$$\chi^u(E)=3,\ \chi^u(C_4^2)=-1,\ \chi^u(C_4)=1,\ \chi^u(JC_4^2)=1,\ \chi^u(JC_2)=1. \quad (8.18)$$

Эти характеры выписаны в последней строке табл. III.3. Разлагая представление D^u_Δ с характерами (8.18) по неприводимым представлениям группы $\boldsymbol{C}_{4v}$, получим

$$D^u_\Delta=\Delta_1+\Delta_5. \quad (8.19)$$

Тот же результат может быть получен, если учесть, что переход из центра бриллюэновской зоны Г на линию Δ связан с уменьшением симметрии от $\boldsymbol{O}_h$ до $\boldsymbol{C}_{4v}$. Такое понижение симметрии, как известно (гл. II, § 7, п. 2), вызовет расщепление уровня Γ_{15}.

Выписывая характеры Γ_{15}, соответствующие классам группы C_{4v}, видим, что они совпадают с характерами χ^a (8.18), поэтому, действуя по общему рецепту, получим

$$\Gamma_{15} = \Delta_1 + \Delta_5, \tag{8.19a}$$

т. е. трехкратно вырожденное в центре состояние Γ_{15} расщепляется вдоль линии Δ на невырожденное состояние Δ_1 и дважды вырожденное состояние Δ_5 [1]).

Однако первый способ, использующий (8.18), имеет то преимущество, что он применим и к точке T (см. рис. III.13), не соприкасающейся с центром Г, но обладающей той же симметрией C_{4v}, что и точка Δ.

Сравнивая (8.19а) с аналитическим решением уравнений механики в предыдущем параграфе (табл. III.1, рис. III.14, *а*), мы видим, что теория групп только на основании соображений, связанных с симметрией кубической решетки, предсказывает существование вдоль линии Δ зоны Бриллюэна двух ветвей колебаний — невырожденной продольной (Δ_1) и дважды вырожденной поперечной (Δ_5). Аналогично для линии Λ получим из табл. III.4

$$\Gamma_{15} = \Lambda_1 + \Lambda_3, \tag{8.20}$$

т. е. трехкратно вырожденное состояние Γ_{15} вдоль линии Λ (см. рис. III.13) расщепляется на единичное представление Λ_1 и дважды вырожденное представление Λ_3.

Мы уже отмечали, что симметрия точки R совпадает с симметрией точки Г, поэтому очевидно, что в точке R состояния Λ_1 и Λ_3 сливаются в трехкратно вырожденное состояние. Тем самым качественно описывается дисперсионная картина на рис. III.14, *б*. Наконец, на линии Σ получим из табл. III.4

$$\Gamma_{15} = \Sigma_1 + \Sigma_3 + \Sigma_4. \tag{8.21}$$

Таким образом, вдоль линии Σ происходит полное расщепление состояния Γ_{15} на три невырожденных состояния, как это следует и из аналитического решения (табл. III.1, в, рис. III.14, *в*).

Мы видим, что состояние Γ_{15} в центре бриллюэновской зоны при смещении вдоль линии Δ, Λ и Σ однозначно разлагается на состояния (8.19а), (8.20) и (8.21). Вообще, если элементы симметрии соприкасаются, то неприводимые представления группы более высокой симметрии однозначно разлагаются на неприводимые представления групп более низкой симметрии. Эти условия носят название *соотношений совместности*. В табл. III.5 представлены соотношения совместности между неприводимыми представлениями в точке Г и точках Δ, Λ и Σ.

---

[1]) Именно поэтому неприводимому представлению Γ_{15} приписывается индекс «15».

Для завершения группового анализа ветвей колебаний простого кубического кристалла надо составить таблицы характеров групп волновых векторов для точек X и M, имеющей ту же симметрию, и Z.

Таблица III.5

Γ_1	Γ_2	Γ_{12}	Γ'_{15}	Γ'_{25}
Δ_1 Λ_1 Σ_1	Δ_2 Λ_2 Σ_4	$\Delta_1\Delta_2$ Λ_3 $\Sigma_1\Sigma_4$	$\Delta'_1\Delta_5$ $\Lambda_2\Lambda_3$ $\Sigma_2\Sigma_3\ \Sigma_4$	$\Delta'_2\Delta_5$ $\Lambda_1\Lambda_3$ $\Sigma_1\Sigma_2\Sigma_3$
Γ'_1	Γ'_2	Γ'_{12}	Γ_{15}	Γ_{25}
Δ'_1 Λ_2 Σ_2	Δ'_2 Λ_1 Σ_3	$\Delta'_1\Delta'_2$ Λ_3 $\Sigma_2\Sigma_3$	$\Delta_1\Delta_5$ $\Lambda_1\Lambda_3$ $\Sigma_1\Sigma_3\Sigma_4$	$\Delta_2\Delta_5$ $\Lambda_2\Lambda_3$ $\Sigma_1\Sigma_2\Sigma_4$

Для определения группы волнового вектора точки X (и M) следует учесть, что точка X приобретает по сравнению с точкой Δ дополнительные элементы симметрии, связанные с инверсией J точки X; при этом мы получаем эквивалентный вектор $\boldsymbol{q}$, отличающийся от исходного на вектор обратной решетки $2\pi/a$. Таким образом, группа волнового вектора $G_{q_X} = G_{q_\Delta} \times C_i$ и ее характеры могут быть получены из табл. III.3 по схеме табл. II.6. Если затем составить таблицы соотношений совместности для M и Σ, Z, T и для X и Δ, Z, S, то можно предсказать характер слияния ветвей в точке M в случае в) и характер расщепления ветвей в случае г).

Рассмотрим в заключение неприводимые представления группы волнового вектора, направленного к точке Z (см. рис. III.13). Поскольку линия Z не выходит из центра Г, мы не можем пользоваться условиями совместности, а должны исходить из представления D^u_Z (8.7). Волновой вектор, направленный к точке Z, не меняется (или переходит в эквивалентный) при следующих операциях: E, $C_4^2(q_z)$, $JC_4^2(q_x)$, $JC_4^2(q_y)$ (в скобках указаны соответствующие оси C_4). Таблица характеров этой группы четвертого порядка совпадает с таблицей характеров группы для точек Σ и S (см. табл. III.4)[1]. Легко показать, что характеры представления D^u_Z совпадают со значениями последней строки для Γ_{15}; поэтому разложение D^u_Z по Z_1, Z_2, ... имеет вид,

[1]) Необходимо только заменить первый столбец на неприводимые представления Z_1, Z_2, Z_3 и Z_4, а первую строку на E, $C_4^2(q_z)$, $JC_4^2(q_x)$, $JC_4^2(q_y)$.

аналогичный (8.21), т. е.

$$D_Z^u = Z_1 + Z_3 + Z_4. \tag{8.22}$$

Мы видим, что вдоль линии Z имеются три невырожденные ветви колебаний, как и было показано аналитически (см. табл. III.1, рис. III.14, *г*).

Учитывая, как сравнительно просто было в предыдущем параграфе получено решение для простой кубической решетки, читатель может усомниться в эффективности применения теории групп к нормальным колебаниям кристалла. В связи с этим отметим, что, во-первых, решение предыдущего параграфа получено для модели кубического кристалла, в котором учитывается взаимодействие каждого атома только с первой и второй координационной группой атомов. В то время как результаты группового анализа справедливы для кубического кристалла в общем случае. Во-вторых, теория групп может быть применена к анализу нормальных колебаний и в случае сложных решеток, когда получить аналитическое решение в замкнутой форме невозможно. При этом теория групп позволяет отбросить все то, что противоречит симметрии системы, позволяя тем самым существенно уменьшить числа вариантов, которые должны быть исследованы методами вычислительной математики.

§ 9. Колебания и волны в кристаллах в приближении изотропного континуума

Рассмотрим гармонические волны в кристаллической решетке в континуальном приближении (П. Дебай, 1912). Очевидно, что такое приближение является хорошим, когда длина волны много больше постоянной решетки, так как в этом случае не должна сказываться дискретная (атомная) структура кристалла. Как мы увидим, континуальное приближение носит весьма различный характер для длинных акустических волн и для длинных оптических волн в ионном (гетерополярном) кристалле.

1. В случае длинных акустических волн континуальное приближение эквивалентно применению теории упругости [1]). Уравнения движения однородного, изотропного, упругого континуума в отсутствие объемных сил имеют вид [2])

$$\rho \frac{\partial^2 \boldsymbol{u}}{\partial t^2} = (M + \Lambda) \operatorname{grad} \operatorname{div} \boldsymbol{u} + M \nabla^2 \boldsymbol{u}. \tag{9.1}$$

Здесь $\boldsymbol{u}(\boldsymbol{r}, t)$ — вектор смещения среды в точке $\boldsymbol{r}$ в момент t,

[1]) При изложении этого пункта я следую § 2 гл. VI моей книги «Основы статистической физики и термодинамики».— М., 1973 г.

[2]) Ландау Л. Д., Лифшиц Е. М./Теоретическая физика, т. 7.— Теория упругости.— 3 изд.— М.: Наука, 1965, § 22.

M и Λ — постоянные коэффициенты Ламэ, ρ — постоянная плотность однородного континуума.

Из теории упругости известно [1]), что $\vartheta \equiv \operatorname{div} \boldsymbol{u}$ — относительное изменение объема $\Delta V/V$ в точке $\boldsymbol{r}$, а $\varphi = {}^1/_2 \operatorname{rot} \boldsymbol{u}$ — угол поворота элемента объема (как целого) в точке $\boldsymbol{r}$. Беря дивергенцию от обеих частей равенства (9.1), получим волновое уравнение для сжатия ϑ

$$\frac{\partial^2 \vartheta}{\partial t^2} = v_l^2 \nabla^2 \vartheta, \tag{9.2}$$

где $v_l = \sqrt{(2M+\Lambda)/\rho}$ — скорость волн сжатия; при получении (9.2) мы использовали соотношения

$$\operatorname{div} \frac{\partial^2 \boldsymbol{u}}{\partial t^2} = \frac{\partial^2}{\partial t^2} (\operatorname{div} \boldsymbol{u}), \quad \operatorname{div} \operatorname{grad} \equiv \nabla^2, \quad \operatorname{div} \nabla^2 \boldsymbol{u} = \nabla^2 \operatorname{div} \boldsymbol{u}.$$

Аналогично, беря ротор обеих частей уравнения (9.1), получим волновое уравнение для угла кручения $\boldsymbol{\varphi}$:

$$\frac{\partial^2 \boldsymbol{\varphi}}{\partial t^2} = v_t \nabla^2 \boldsymbol{\varphi}, \tag{9.3}$$

где $v_t = \sqrt{M/\rho}$ — скорость распространения волн кручения (было использовано, что $\operatorname{rot} \operatorname{grad} \equiv 0$). Легко видеть, что $v_l > v_t$, что связано с тем, что упругое сопротивление при сжатии больше, чем при кручении.

Покажем, что волны сжатия продольные, а волны кручения поперечные. Рассмотрим плоскую волну, распространяющуюся вдоль оси x (что, очевидно, не является ограничением общности):

$$\boldsymbol{u} = \boldsymbol{A} \sin 2\pi \left(\nu t - \frac{x}{\lambda} \right), \tag{9.4}$$

где $\boldsymbol{A}$ — постоянная амплитуда, ν — частота и λ — длина волны. Отсюда следует, что

$$\vartheta = \operatorname{div} \boldsymbol{u} = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z} = -A_x \left(\frac{2\pi}{\lambda} \right) \cos 2\pi \left(\nu t - \frac{x}{\lambda} \right) \tag{9.5}$$

и

$$\boldsymbol{\varphi} = \frac{1}{2} \operatorname{rot} \boldsymbol{u} = -A_y \boldsymbol{j}_0 \left(\frac{\pi}{\lambda} \right) \cos 2\pi \left(\nu t - \frac{x}{\lambda} \right) + \\ + A_z \boldsymbol{k}_0 \left(\frac{\pi}{\lambda} \right) \cos 2\pi \left(\nu t - \frac{x}{\lambda} \right), \tag{9.6}$$

где $\boldsymbol{j}_0$ и $\boldsymbol{k}_0$ — орты осей y и z.

Из (9.5) и (9.6) непосредственно видно, что волны сжатия ϑ являются продольными ($A_y = A_z = 0$), а волны кручения φ — поперечными ($A_x = 0$).

[1]) Ландау Л. Д., Лифшиц Е. М. Теория упругости.— 3 изд.— М.: Наука, 1965, § 22.

Продольные волны сжатия ϑ (9.5) и поперечные волны кручения φ_y и φ_z (9.6) являются континуальным аналогом трех акустических ветвей, рассмотренных при изучении колебаний атомов сложных кристаллов, в § 5, п. 5. Можно сказать, что v_l и v_t — продольная и поперечная скорости звука. Определим в рассматриваемом нами случае функцию распределения частот $g(\omega)$ (5.33), дающую число колебаний на единичный интервал частоты ω.

Волновое уравнение для скаляра ϑ (9.2) аналогично волновому уравнению (9.3) для каждой составляющей вектора кручения $\boldsymbol{\varphi}$, поэтому достаточно рассмотреть уравнение (9.2). Рассмотрим волны ϑ в кубе с ребром L. Направим прямоугольные оси координат x, y и z по ребрам куба; выберем в качестве граничных условий: $\vartheta = 0$ на всех шести гранях куба $x = y = z = 0$ и $x = y = z = L$. Выбор граничных условий не может быть существен, если длина волны мала по сравнению с L. В частности, ничего не изменилось бы, если бы мы потребовали равенства нулю на гранях куба не сжатия ϑ, а упругих напряжений (свободные поверхности).

Ищем решение (9.2) в виде

$$\vartheta = A \sin \omega t \sin ax \sin by \sin cz, \tag{9.7}$$

где A — амплитуда, ω — циклическая частота колебаний и a, b, c — постоянные. Подставляя (9.7) в (9.2), получим после сокращения обеих частей на ϑ:

$$\omega = v_l \sqrt{a^2 + b^2 + c^2}. \tag{9.8}$$

Таким образом, выражение (9.7) удовлетворяет уравнению (9.2), если частота ω связана с a, b, c соотношением (9.8). Для того чтобы удовлетворить граничным условиям, мы должны положить

$$aL = n_1\pi, \quad bL = n_2\pi, \quad cL = n_3\pi, \tag{9.9}$$

где n_1, n_2, n_3 — целые положительные числа или нуль (отрицательные числа приводят к тому же колебанию с фазой, сдвинутой на π).

Подставляя (9.9) в (9.8), получим

$$\omega = \frac{\pi v_l}{L} \sqrt{n_1^2 + n_2^2 + n_3^2}. \tag{9.10}$$

Каждой тройке чисел n_i соответствует свое нормальное колебание с определенной частотой (9.10). Если n_1, n_2, n_3 — большие числа, т. е. длина волны колебаний много меньше L, то ω зависит от чисел n_i квазинепрерывно. В этом случае можно поставить вопрос о числе колебаний в интервале частоты $(\omega, \omega + d\omega)$.

Введем величину

$$R^2 = n_1^2 + n_2^2 + n_3^2, \tag{9.11}$$

тогда частота

$$\omega = \frac{\pi v_l}{L} R. \tag{9.12}$$

На рис. III.15 представлена декартова координатная система, по осям которой отложены целые положительные числа n_1, n_2, n_3; каждому узлу кубической решетки, изображенной на рисунке [1]), соответствуют целочисленные координаты n_1, n_2, n_3 и определенный радиус-вектор R (9.11). С другой стороны, каждому узлу решетки соответствует определенное нормальное колебание (9.7) с частотой (9.10).

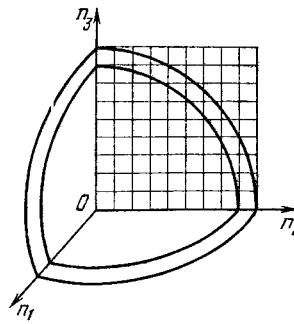


Рис. III.15.

Определим число колебаний на интервал частоты $(\omega, \omega + d\omega)$, когда числа n_l велики. Из (9.12) и (9.11) следует, что это число равно числу узлов решетки в шаровом слое $(R, R + dR)$ в координатном октанте. Поскольку объем кубической ячейки равен единице, это число просто равно объему соответствующего шарового слоя.

Таким образом, число продольных колебаний на интервал частоты $(\omega, \omega + d\omega)$, равно

$$g_l(\omega)\, d\omega = \frac{4\pi R^2\, dR}{8} = \frac{V}{2\pi^2 v_l^3}\, \omega^2 d\omega, \tag{9.13}$$

как это следует из (9.12); здесь $V = L^3$ — объем тела. Для каждой из двух составляющих вектора φ, удовлетворяющих уравнению (9.3), справедливы аналогичные соображения; поэтому число поперечных колебаний в интервале $(\omega, \omega + d\omega)$ равно

$$g_t(\omega)\, d\omega = \frac{2V}{2\pi^2 v_t^3}\, \omega^2\, d\omega, \tag{9.14}$$

где два в числителе учитывает наличие двух составляющих для поперечной волны (9.6).

Полная функция распределения частот (5.32) равна

$$g(\omega) = g_l(\omega) + g_t(\omega) = \frac{3V}{2\pi^2 v_0^3}\, \omega^2, \tag{9.15}$$

где скорость v_0 определяется из равенства

$$\frac{1}{v_0^3} = \frac{1}{3}\left(\frac{1}{v_l^3} + \frac{2}{v_t^3}\right). \tag{9.16}$$

[1]) Для того чтобы не усложнять рисунка, на нем представлены узлы решетки только в плоскости (n_2, n_3).

Можно сказать, что v_0^3 есть средняя от обратного куба продольной и поперечной скорости звука.

Отметим, что в континуальном (дебаевском) приближении для акустических колебаний функция распределения частот $g(\omega)$ (5.32) пропорциональна ω^2.

2. В предыдущем пункте мы видели, как в континуальном приближении описываются длинные акустические волны.

Исследуем теперь континуальное приближение для длинных оптических волн в кубическом ионном кристалле (К. Хуанг, 1950). Рассмотрим ионный кристалл, каждая ячейка которого состоит из двух разноименных ионов с эффективными зарядами $\pm e^*$ и массами m_+ и m_-. В длинноволновых оптических колебаниях движение ионов во всех ячейках происходит синхронно (§ 5, п. 5), поэтому достаточно рассмотреть движение ионов в одной ячейке. Если $\boldsymbol{u}_+$ и $\boldsymbol{u}_-$ — смещения положительного и отрицательного ионов из положений равновесия, то уравнения движения имеют вид

$$m_+ \frac{d^2\boldsymbol{u}_+}{dt^2} = -\varkappa(\boldsymbol{u}_+ - \boldsymbol{u}_-) + e^*\boldsymbol{E}_e, \tag{9.17}$$

$$m_- \frac{d^2\boldsymbol{u}_-}{dt^2} = -\varkappa(\boldsymbol{u}_- - \boldsymbol{u}_+) - e^*\boldsymbol{E}_e. \tag{9.17а}$$

Здесь $\boldsymbol{E}_e$ — эффективное электрическое поле, действующее на ион со стороны внешнего поля и остальных ионов кристалла, $\varkappa$ — коэффициент квазиупругой силы, действующей на ион при смещении его относительно другого иона. Деля (9.17) на m_+, (9.17а) на m_- и вычитая их почленно, получим

$$m_r \frac{d^2\boldsymbol{s}}{dt^2} = -\varkappa\boldsymbol{s} + e^*\boldsymbol{E}_e, \tag{9.18}$$

где m_r — приведенная масса ($m_r^{-1} = m_+^{-1} + m_-^{-1}$), $\boldsymbol{s} = \boldsymbol{u}_+ - \boldsymbol{u}_-$ — смещение положительного иона относительно отрицательного иона.

Эффективное поле в кубическом ионном кристалле [1])

$$\boldsymbol{E}_e = \boldsymbol{E} + \frac{4\pi}{3}\boldsymbol{P}, \tag{9.19}$$

где $\boldsymbol{E}$ — среднее поле в диэлектрике, а вектор поляризации

$$\boldsymbol{P} = N_0[e^*(\boldsymbol{u}_+ - \boldsymbol{u}_-) + \alpha_+\boldsymbol{E}_e + \alpha_-\boldsymbol{E}_e] = N_0[e^*\boldsymbol{s} + \alpha\boldsymbol{E}_e]. \tag{9.20}$$

Здесь N_0 — число ячеек в единице объема кристалла, α_+ и α_- — электронные поляризуемости положительных и отрицательных ионов, а $\alpha = \alpha_+ + \alpha_-$. Континуальность подхода связана, в частности, с использованием макроскопического понятия вектора поляризации $\boldsymbol{P}$.

---

[1]) Тамм И. Е. Основы теории электричества. — 9 изд. — М.: Наука, 1976, § 28.

Исключая из (9.19) и (9.20) эффективное поле $\boldsymbol{E}_e$, получим

$$\boldsymbol{P}=N_0\frac{e^*\boldsymbol{s}+\alpha E}{1-\frac{4\pi N_0}{3}\alpha}. \tag{9.21}$$

Для того чтобы исключить из формул неизмеряемую непосредственно поляризуемость α, будем исходить из выражения для вектора индукции [1])

$$\boldsymbol{D}=\boldsymbol{E}+4\pi\boldsymbol{P}=\varepsilon\boldsymbol{E}, \tag{9.22}$$

где ε — диэлектрическая постоянная [2]); отсюда

$$\boldsymbol{P}=\frac{\varepsilon-1}{4\pi}\boldsymbol{E}. \tag{9.23}$$

В высокочастотном ($\omega\to\infty$) электрическом поле ионы не успевают следовать за его изменением, поэтому $\boldsymbol{s}\to 0$. Исключим для этого случая $\boldsymbol{P}$ из (9.21) (где мы положим $\boldsymbol{s}=0$) и (9.23) (где мы положим $\varepsilon=\varepsilon_\infty$); тогда поляризуемость

$$\alpha=\frac{\varepsilon_\infty-1}{\frac{4\pi N_0}{3}(\varepsilon_\infty+2)}. \tag{9.24}$$

Подставляя это значение α в (9.21), получим

$$\boldsymbol{P}=N_0\frac{e^*(\varepsilon_\infty+2)}{3}\boldsymbol{s}+\frac{\varepsilon_\infty-1}{4\pi}\boldsymbol{E}. \tag{9.25}$$

Подставляя это значение $\boldsymbol{P}$ в (9.19), которое затем используется для исключения $\boldsymbol{E}_e$ из (9.18), получим

$$m_r\frac{d^2\boldsymbol{s}}{dt^2}=-m_r\omega_0^2\boldsymbol{s}+\frac{e^*(\varepsilon_\infty+2)}{3}\boldsymbol{E}, \tag{9.26}$$

где

$$\omega_0^2=\frac{\varkappa}{m_r}-\frac{4\pi N_0 e^{*2}(\varepsilon_\infty+2)}{9m_r}. \tag{9.26а}$$

Введем «нормированное» отклонение

$$\boldsymbol{w}=\sqrt{N_0 m_r}\,\boldsymbol{s} \tag{9.27}$$

и подставим его в (9.26) и (9.25); тогда

$$\frac{d^2\boldsymbol{w}}{dt^2}=-\omega_0^2\boldsymbol{w}+\sqrt{\frac{N_0}{m_r}}\,e^*\frac{\varepsilon_\infty+2}{3}\boldsymbol{E} \tag{9.28}$$

и

$$\boldsymbol{P}=\sqrt{\frac{N_0}{m_r}}\,e^*\frac{\varepsilon_\infty+2}{3}\boldsymbol{w}+\frac{\varepsilon_\infty-1}{4\pi}\boldsymbol{E}. \tag{9.29}$$

---

[1]) Тамм И. Е. Основы теории электричества.— 9 изд.— М.: Наука, 1976, § 22.

[2]) Мы используем здесь привычное обозначение для диэлектрической постоянной, не опасаясь, что ε будет спутано с энергией.

Если ввести статическую диэлектрическую постоянную ε_0 ($\omega \to 0$) в постоянном электрическом поле, то из (9.23) и (9.29) следует

$$\varepsilon_0 - \varepsilon_\infty = \frac{N_0}{m_r} e^{*2} \frac{4\pi (\varepsilon_\infty + 2)^2}{9\omega_0^2}. \tag{9.30}$$

Определим отсюда величину $(N_0/m_r)\, e^{*2}$ и подставим ее в (9.28) и (9.29), тогда получим

$$\frac{d^2 \boldsymbol{w}}{dt^2} = -\omega_0^2 \boldsymbol{w} + \omega_0 \sqrt{\frac{\varepsilon_0 - \varepsilon_\infty}{4\pi}}\, \boldsymbol{E} \tag{9.31}$$

и

$$\boldsymbol{P} = \omega_0 \sqrt{\frac{\varepsilon_0 - \varepsilon_\infty}{4\pi}}\, \boldsymbol{w} + \frac{\varepsilon_\infty - 1}{4\pi} \boldsymbol{E}. \tag{9.32}$$

Для определения характера движения ионов мы положим

$$\boldsymbol{w} = \boldsymbol{w}_t + \boldsymbol{w}_l, \tag{9.33}$$

где

$$\operatorname{div} \boldsymbol{w}_t = 0, \quad \operatorname{rot} \boldsymbol{w}_l = 0. \tag{9.34}$$

Известно, что такое представление любого вектора ($\boldsymbol{w}$) в виде суммы соленоидального вектора ($\boldsymbol{w}_t$) и потенциального вектора ($\boldsymbol{w}_l$) всегда возможно и единственно [1]).

В отсутствие свободных зарядов

$$\operatorname{div} \boldsymbol{D} = \operatorname{div} \boldsymbol{E} + 4\pi \operatorname{div} \boldsymbol{P} = 0. \tag{9.35}$$

Если подставить сюда $\boldsymbol{P}$ из (9.32) и воспользоваться (9.33) и (9.34), то получим

$$\operatorname{div} \boldsymbol{E} + \frac{\omega_0}{\varepsilon_\infty} \sqrt{4\pi (\varepsilon_0 - \varepsilon_\infty)}\, \operatorname{div} \boldsymbol{w}_l = 0, \tag{9.36}$$

откуда

$$\boldsymbol{E} = -\frac{\omega_0}{\varepsilon_\infty} \sqrt{4\pi (\varepsilon_0 - \varepsilon_\infty)}\, \boldsymbol{w}_l. \tag{9.37}$$

Подставляя (9.33) и (9.37) в (9.31), получим

$$\frac{d^2}{dt^2} (\boldsymbol{w}_t + \boldsymbol{w}_l) = -\omega_0^2 \boldsymbol{w}_t - \omega_0^2 \frac{\varepsilon_0}{\varepsilon_\infty} \boldsymbol{w}_l. \tag{9.38}$$

Разделяя в этом уравнении соленоидальную и потенциальную части, получим

$$\frac{d^2 \boldsymbol{w}_t}{dt^2} = -\omega_0^2 \boldsymbol{w}_t, \tag{9.39}$$

$$\frac{d^2 \boldsymbol{w}_l}{dt^2} = -\omega_0^2 \frac{\varepsilon_0}{\varepsilon_\infty} \boldsymbol{w}_l. \tag{9.40}$$

Если мы представим $\boldsymbol{w}_t$ или $\boldsymbol{w}_l$ в виде плоской волны $\boldsymbol{A} \exp[i(\boldsymbol{qr} - \omega t)]$, то получим из (9.39) и (9.40) для соответствующих частот: $\omega_t = \omega_0$ и $\omega_l = (\varepsilon_0/\varepsilon_\infty)^{1/2} \omega_0$.

[1]) Кочин Н. Е. Векторное исчисление и начало тензорного исчисления.— 9 изд.— М.: Наука, 1965.

С другой стороны, подставляя в (9.34) выражения для плоской волны, получим

$$\begin{aligned} \operatorname{div} \boldsymbol{w}_t &= \operatorname{div}\{\boldsymbol{A}_t \exp[i(\boldsymbol{qr}-\omega_t t)]\} \backsim \boldsymbol{A}_t \boldsymbol{q} = 0, \\ \operatorname{rot} \boldsymbol{w}_t &= \operatorname{rot}\{\boldsymbol{A}_t \exp[i(\boldsymbol{qr}-\omega_t t)]\} \backsim [\boldsymbol{A}_t \boldsymbol{q}] = 0, \end{aligned} \tag{9.41}$$

откуда следует, что $\boldsymbol{A}_t \perp \boldsymbol{q}$ и $\boldsymbol{A}_l \parallel \boldsymbol{q}$, т. е. соленоидальная волна $\boldsymbol{w}_t$ поперечная, а потенциальная волна $\boldsymbol{w}_l$ продольная.

Из полученных выше выражений следует соотношение Лиддена—Сакса—Теллера

$$\omega_l/\omega_t = \sqrt{\varepsilon_0/\varepsilon_\infty}. \tag{9.42}$$

Так как $\varepsilon_0 > \varepsilon_\infty$ (см. (9.30)), то частота продольных волн ω_l больше частоты поперечных волн ω_t, что аналогично соотношению $v_l > v_t$ для акустических волн (см. предыдущий пункт). Так как измерить на опыте ω_t проще, чем ω_l, то формула (9.42) может служить для определения ω_l.

§ 10. Квантование колебаний кристаллической решетки. Фононы

1. Квантовомеханический гамильтониан в Q-представлении, соответствующий функции Гамильтона нормальных колебаний решетки (6.21), получается из нее при замене импульсов $\boldsymbol{P}_j(\boldsymbol{q})$ операторами

$$\dot{Q}_j(\boldsymbol{q}) = P_j(\boldsymbol{q}) \longrightarrow \frac{\hbar}{i}\frac{\partial}{\partial Q_j(\boldsymbol{q})}, \tag{10.1}$$

где $\hbar$—постоянная Планка, деленная на 2π, $i = \sqrt{-1}$.

Гамильтониан, следующий из (6.21), имеет вид

$$\hat{\mathscr{H}}(Q, P) = \sum_{j,\,\boldsymbol{q}} \left\{ -\frac{\hbar^2}{2}\frac{\partial^2}{\partial Q_j^2(\boldsymbol{q})} + \frac{1}{2}\omega_j^2(\boldsymbol{q})\, Q_j^2(\boldsymbol{q}) \right\}, \tag{10.2}$$

т. е. распадается на сумму, каждое слагаемое которой имеет вид гамильтониана линейного гармонического осциллятора с координатой $Q_j(\boldsymbol{q})$, частотой $\omega_j(\boldsymbol{q})$ и массой, равной единице. Если гамильтониан системы состоит из суммы, каждое слагаемое которой зависит только от одной координаты и оспряженного ей импульса, то, как известно из квантовой механики, волновая функция системы равна произведению волновых функций, соответствующих каждому слагаемому, а энергия равна сумме соответствующих энергий.

Рассмотрим отдельное слагаемое гамильтониана (10.2), причем для простоты записи спустим символы j и $\boldsymbol{q}$ у координаты $Q_j(\boldsymbol{q})$; тогда уравнение Шредингера, соответствующее такому линейному

осциллятору, имеет вид

$$-\frac{\hbar^2}{2}\frac{\partial^2\psi}{\partial Q^2}+\frac{1}{2}\omega^2Q^2\psi=\varepsilon\psi. \tag{10.3}$$

Из квантовой механики известно[1]), что собственные значения и нормированные собственные функции этого уравнения имеют вид

$$\varepsilon\equiv\varepsilon_N=\hbar\omega\,(N+{}^1/_2), \tag{10.4а}$$

$$\psi\equiv\psi_N(Q)=\left(\frac{\omega}{\pi\hbar}\right)^{1/4}\frac{1}{\sqrt{2^N N!}}\,e^{-\omega Q^2/2\hbar}H_N\left[\left(\frac{\omega}{\hbar}\right)^{1/2}Q\right]. \tag{10.4б}$$

Здесь N — квантовое число осциллятора, $H_N[\xi]$ — полином Эрмита от безразмерной координаты $\xi=\sqrt{(1\text{г}\times\omega)/\hbar}\,Q=(\omega/\hbar)^{1/2}Q$.

Как мы увидим дальше, при изучении взаимодействия электронов проводимости с колебаниями решетки существенное значение имеют матричные элементы координаты Q и импульса P (10.1) на волновых функциях (10.4б). Можно показать[2]), что они равны

$$\langle N'|Q|N\rangle=\int\limits_{-\infty}^{\infty}\psi_{N'}^{*}\,Q\psi_N\,dQ=$$

$$=\sqrt{\frac{\hbar}{2\omega}}\times\begin{cases}\sqrt{N}, & \text{если } N'=N-1,\\ \sqrt{N+1}, & \text{если } N'=N+1,\\ 0 & \text{во всех остальных случаях;}\end{cases} \tag{10.5}$$

$$\langle N'|\hat{P}|N\rangle=\frac{\hbar}{i}\int\limits_{-\infty}^{\infty}\psi_{N'}^{*}\,\frac{\partial\psi_N}{\partial Q}\,dQ=$$

$$=i\sqrt{\frac{\hbar\omega}{2}}\times\begin{cases}-\sqrt{N}, & \text{если } N'=N-1,\\ \sqrt{N+1}, & \text{если } N'=N+1,\\ 0 & \text{во всех остальных случаях.}\end{cases} \tag{10.6}$$

2. Введем вместо $a_j(\boldsymbol{q})$ новые комплексные нормальные координаты $\mathfrak{a}_j(\boldsymbol{q})$, полагая

$$a_j(\boldsymbol{q})=\sqrt{\frac{\hbar}{2\omega_j(\boldsymbol{q})}}\left[\mathfrak{a}_j^{*}(-\boldsymbol{q})+\mathfrak{a}_j(\boldsymbol{q})\right], \tag{10.7}$$

$$\dot{a}_j(\boldsymbol{q})=i\sqrt{\frac{\hbar\omega_j(\boldsymbol{q})}{2}}\left[\mathfrak{a}_j^{*}(-\boldsymbol{q})-\mathfrak{a}_j(\boldsymbol{q})\right]. \tag{10.8}$$

---

[1]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика.— 3 изд.— М., 1974, § 23.

[2]) Там же.

Определенные таким образом $a_j(\boldsymbol{q})$ автоматически удовлетворяют условию (6.5). Решая систему (10.7), (10.8), относительно $\mathfrak{a}_j(\boldsymbol{q})$ и $\mathfrak{a}_j(-\boldsymbol{q})$, получим

$$\mathfrak{a}_j(\boldsymbol{q}) = \sqrt{\frac{\omega_j(\boldsymbol{q})}{2\hbar}}\, a_j(\boldsymbol{q}) + i\sqrt{\frac{1}{2\hbar\omega_j(\boldsymbol{q})}}\, \dot{a}_j(\boldsymbol{q}). \tag{10.9}$$

Выражая $a_j(\boldsymbol{q})$ и $\dot{a}_j(\boldsymbol{q})$, посредством (6.16), через $Q_j(\boldsymbol{q})$ и $\dot{Q}_j(\boldsymbol{q}) = P_j(\boldsymbol{q})$, получим

$$\mathfrak{a}_j(\boldsymbol{q}) = \sqrt{\frac{\omega_j(\boldsymbol{q})}{2\hbar}}\, Q_j(\boldsymbol{q}) + i\sqrt{\frac{1}{2\hbar\omega_j(\boldsymbol{q})}}\, P_j(\boldsymbol{q}). \tag{10.10}$$

3. Введем операторы,[1] соответствующие величинам $\mathfrak{a}_j(\boldsymbol{q})$ и $\mathfrak{a}_j^*(\boldsymbol{q})$. Опуская для простоты записи индекс j и аргумент $\boldsymbol{q}$, получим из (10.10) и (10.1)

$$a = \left(\frac{\omega}{2\hbar}\right)^{1/2} Q + \left(\frac{\hbar}{2\omega}\right)^{1/2} \frac{\partial}{\partial Q}, \tag{10.11}$$

$$a^+ = \left(\frac{\omega}{2\hbar}\right)^{1/2} Q - \left(\frac{\hbar}{2\omega}\right)^{1/2} \frac{\partial}{\partial Q} \tag{10.12}$$

(a^+ — оператор величины $\mathfrak{a}_j^*(\boldsymbol{q})$).

Легко показать, что эти операторы не эрмитовы. Применяя их к волновой функции (10.4б), получим

$$a\psi_N = \sqrt{N}\,\psi_{N-1}, \tag{10.13}$$

$$a^+\psi_N = \sqrt{N+1}\,\psi_{N+1}. \tag{10.14}$$

В самом деле, вычисляя

$$a\psi_N = \left[\left(\frac{\omega}{2\hbar}\right)^{1/2} Q + \left(\frac{\hbar}{2\omega}\right)^{1/2} \frac{\partial}{\partial Q}\right] \frac{\omega^{1/4}\exp(-\omega Q^2/2\hbar)\, H_N\left[\left(\frac{\omega}{\hbar}\right)^{1/2} Q\right]}{(\pi\hbar)^{1/4}(2^N N!)^{1/2}},$$

получим (10.13); при этом произведение первого слагаемого в квадратной скобке на волновую функцию ψ_N сокращается с производной $\partial/\partial Q$ от экспоненциального множителя в ψ_N; кроме того, надо воспользоваться тем, что[1])

$$H_N'(\xi) = 2NH_{N-1}(\xi). \tag{10.15}$$

Аналогично получается (10.14).

Как мы увидим дальше, при взаимодействии нормальных колебаний решетки с электронами проводимости и взаимодействии колебаний друг с другом эти колебания ведут себя как

---

[1]) Смирнов В. И., т. III, ч. 2, § 158.

частицы с энергией $\hbar\omega_j(\boldsymbol{q})$ и *квазиимпульсом*[1]) $\hbar\boldsymbol{q}$. Эти *квазичастицы* получили название *фононов*.

В Q-представлении состояние кристалла с гамильтонианом (10.2) задается симметризованным произведением осцилляторных функций (10.4б)[2]), каждая из которых характеризуется квантовым числом $N \equiv N_{j\boldsymbol{q}}$. В силу квантовой неразличимости фононов одного сорта $(j, \boldsymbol{q})$ состояние кристалла полностью описывается числами фононов $N_{j\boldsymbol{q}}$. Такой способ описания систем квантовотождественных частиц получил название *метода вторичного квантования*. Операторы a и a^+, введенные выше, являются операторами в представлении вторичного квантования, так как они непосредственно действуют на *числа заполнения* N. В самом деле, оператор a, как видно из (10.13), уменьшает число фононов N на единицу (поэтому он называется *оператором уничтожения*), а оператор a^+, как видно из (10.14), увеличивает число фононов N на единицу (поэтому он называется *оператором рождения*).

Рассмотрим некоторые свойства операторов a и a^+. Из квантовой механики известно, что для сопряженных координаты $Q \equiv Q_j(\boldsymbol{q})$ и импульса $P \equiv P_j(\boldsymbol{q})$ имеют место перестановочные соотношения[3]), т. е.

$$QP - PQ = [Q, P] = i\hbar. \tag{10.16}$$

Подставляя сюда вместо операторов Q и $P = \dot{Q}$ их выражения через a и a^+ (см. (10.11), (10.12)), получим

$$aa^+ - a^+a = [a, a^+] = 1. \tag{10.17}$$

Так как операторы $Q_j(\boldsymbol{q})$ и $P_{j'}(\boldsymbol{q}')$ при $j \neq j'$ или $\boldsymbol{q} \neq \boldsymbol{q}'$ коммутируют друг с другом, то (10.17) может быть записано в виде

$$[a_j(\boldsymbol{q}), a_{j'}^+(\boldsymbol{q}')] = \delta_{jj'}\delta_{\boldsymbol{q}\boldsymbol{q}'}. \tag{10.18}$$

Квазичастицы (частицы), операторы рождения и уничтожения которых удовлетворяют правилам коммутации (10.18) называются *бозе-частицами* или *бозонами*. В состоянии термодинамического равновесия они описываются *статистикой Бозе—Эйнштейна*[4]).

Выразим гамильтониан (10.2) через операторы a и a^+; воспользовавшись (10.11) и (10.12), получим

$$\mathscr{H}(a) = \sum_{j,\,\boldsymbol{q}} \frac{\hbar\omega}{2}[aa^+ + a^+a], \tag{10.19}$$

---

[1]) Мы называем величину $\hbar\boldsymbol{q}$ квазиимпульсом, а не импульсом, так как она обладает некоторыми особенностями, не свойственными импульсу свободной частицы; например, при столкновении сумма квазиимпульсов сохраняется с точностью до произвольного вектора обратной решетки.

[2]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика.— 3 изд.— М., 1974, § 61.

[3]) Там же.

[4]) Ансельм А. И. Основы статистической физики и термодинамики.— М., 1973.

что можно, используя (10.18), переписать в виде

$$\mathscr{H}(a)=\sum_{j,\,q}\hbar\omega_j\,[a^+a+{}^1/_2]. \tag{10.20}$$

Применяя оператор a^+a к волновой функции (10.4б), получим, учитывая (10.13) и (10.14):

$$a^+a\psi_N=a^+\sqrt{N}\psi_{N-1}=N\psi_N. \tag{10.21}$$

Мы видим, что собственные значения оператора a^+a равны числу частиц (фононов) N_{jq} в состоянии $(j,\,q)$.

Применяя оператор Гамильтона в форме (10.20) к собственной функции ψ_N, получим, используя (10.21), для собственных значений энергии

$$\mathscr{E}=\sum_{j,\,q}\hbar\omega_j(q)\,[N_{jq}+{}^1/_2], \tag{10.22}$$

что согласуется с (10.4а).

Используя (10.13), получим для матричных элементов операторов уничтожения и рождения частиц:

$$\langle N'\,|\,a\,|\,N\rangle=\begin{cases}\sqrt{N}, & \text{если } N'=N-1,\\ 0 & \text{во всех остальных случаях;}\end{cases} \tag{10.23}$$

$$\langle N'\,|\,a^+\,|\,N\rangle=\begin{cases}\sqrt{N+1}, & \text{если } N'=N+1,\\ 0 & \text{во всех остальных случаях.}\end{cases} \tag{10.24}$$

Матричные элементы операторов, соответствующих величинам $a_j(q)$, как следует из (10.7), равны

$$\langle N'_{qj}\,|\,a_j(q)\,|\,N_{qj}\rangle=$$
$$=\begin{cases}\sqrt{\hbar N_{qj}/2\omega_j(q)}, & \text{если } N'_{qj}=N_{qj}-1,\\ 0 & \text{во всех остальных случаях;}\end{cases} \tag{10.25}$$

$$\langle N'_{qj}\,|\,a_j^*(q)\,|\,N_{qj}\rangle=$$
$$=\begin{cases}\sqrt{\hbar\,(N_{qj}+1)/2\omega_j(q)}, & \text{если } N_{qj}=N_{qj}+1,\\ 0 & \text{во всех остальных случаях.}\end{cases} \tag{10.26}$$

При этом мы учитываем, что матричные элементы от операторов $a_j^*(-q)$ при $N'_{qj}=N_{qj}\mp 1$ равны нулю; в самом деле, в Q-представлении (10.10) эти операторы зависят от $Q_j(-q)$, в то время как волновые функции зависят от $Q_j(q)$.

Метод вторичного квантования особенно удобен для квантового рассмотрения систем с очень большим (и даже нефиксированным) числом тождественных частиц (фононов, фотонов, электронов и дырок проводимости в кристалле, электронов и позитронов).

§ 11. Теория теплоемкости кристаллической решетки

Одним из важных применений теории колебаний кристаллических решеток является теория их теплоемкости.

1. Очень проста теория теплоемкости кристаллов в классической области, когда движение атомов решетки подчиняется законам классической механики. Как было показано Л. Больцманом, для систем, находящихся в состоянии статистического равновесия, средняя кинетическая энергия, приходящаяся на одну степень свободы, равна

$$\bar{\varepsilon}_{\text{кин}} = \frac{1}{2}\frac{R}{N_0}T = \frac{1}{2}k_0T, \tag{11.1}$$

где $R = 1{,}986$ *кал/град·моль*—газовая постоянная, $N_0 = 6{,}023 \cdot 10^{23}$—число Авогадро (число частиц в одном моле вещества), $k_0 = R/N_0 = 1{,}38 \cdot 10^{-16}$ *эрг/град*—постоянная Больцмана и T—абсолютная температура [1]).

Легко показать, что для линейного гармонического осциллятора средняя потенциальная энергия $\bar{\varepsilon}_{\text{пот}}$ равна средней кинетической энергии $\bar{\varepsilon}_{\text{кин}}$, поэтому средняя полная энергия

$$\bar{\varepsilon} = \bar{\varepsilon}_{\text{кин}} + \bar{\varepsilon}_{\text{пот}} = 2\bar{\varepsilon}_{\text{кин}}. \tag{11.2}$$

Для одной граммолекулы кристалла элемента его внутренняя энергия в состоянии статистического равновесия, равна энергии $3N_0$ нормальных колебаний

$$\mathcal{E} = 3N_0\bar{\varepsilon} = 3N_0 \cdot 2\bar{\varepsilon}_{\text{кин}} = 3N_0\frac{R}{N_0}T = 3RT. \tag{11.3}$$

Грамм-молекулярная теплоемкость такого кристалла при постоянном объеме равна

$$c_V = \partial\mathcal{E}/\partial T = 3R = 5{,}96 \text{ } \textit{кал/град·моль}, \tag{11.4}$$

т. е. не зависит от температуры и равно примерно 6 *кал/град·моль* (*закон Дюлонга и Пти*).

Большая простота классической теории теплоемкости кристаллов связана с двумя обстоятельствами; с возможностью представления движения атомов кристалла (в гармоническом приближении) в виде нормальных колебаний и универсальностью закона равнораспределения энергии по степеням свободы (11.1). Соотношение (11.4) довольно хорошо оправдывается на опыте, в частности и для металлов, что на первый взгляд не должно иметь места. В самом деле, в металлах число свободных электронов порядка числа атомов; каждый свободный электрон с классической точки зрения обладает средней кинетической энергией $^3/_2k_0T$, что вносит

---

[1]) Ансельм А. И. Основы статистической физики и термодинамики.—М., 1973, гл. V, § 3.

в теплоемкость (11.4) дополнительный вклад $\frac{\partial}{\partial T}\left(N_0 \cdot \frac{3}{2} kT\right) = \frac{3}{2} R$, т. е. увеличивает c_V в полтора раза. Это одно из основных противоречий классической электронной теории металлов будет обсуждено в гл. VI, § 3.

2. Изучение теплоемкости твердых тел при низких температурах показывает, что закон Дюлонга и Пти (11.4) является асимптотическим законом, справедливым только при высоких температурах. При понижении температуры, начиная с некоторой характеристической температуры T_c (так называемой *температуры Дебая*), теплоемкость начинает быстро уменьшаться и при $T \to 0$, $c_V \to 0$. Температура Дебая T_c различна для разных веществ, но для большинства твердых тел она порядка 100—400 °K. Неприменимость закона Дюлонга и Пти к твердым телам при низких температурах обусловлена тем, что при понижении температуры уменьшаются средние скорости движения атомов кристалла и, следовательно, увеличивается соответствующая им дебройлевская длина волны. Из квантовой механики известно, что когда дебройлевская длина волны частицы становится сравнимой или превышает линейные размеры эффективной области движения, движение не подчиняется законам классической механики. По-этому, трактуя нормальные колебания кристалла с квантовой точки зрения, мы должны приписать им, согласно (10.4), энергию

$$\varepsilon_N = \hbar\omega\,(N + {}^1/_2), \tag{11.5}$$

где мы опустили индексы $\boldsymbol{q}$ и j, характеризующие выбранный осциллятор.

Вероятность того, что в состоянии статистического равновесия осциллятор находится в N-м квантовом состоянии с энергией ε_N, равна по Больцману, $w_N = ce^{-\varepsilon_N/k_0T}$. Постоянная c определяется из условия нормировки $\sum\limits_N w_N = c\sum\limits_N e^{-\varepsilon_N/k_0T} = 1$, откуда $c = \left(\sum\limits_N e^{-\varepsilon_N/k_0T}\right)^{-1}$. Средняя энергия осциллятора равна сумме энергий ε_N, помноженных на соответствующие вероятности w_N, т. е.

$$\bar{\varepsilon} = \sum_{N=0}^{\infty} \varepsilon_N w_N = c \sum_{N=0}^{\infty} \varepsilon_N e^{-\varepsilon_N/k_0T} = \frac{\sum\limits_{N=0}^{\infty} \varepsilon_N e^{-\varepsilon_N/k_0T}}{\sum\limits_{N=0}^{\infty} e^{-\varepsilon_N/k_0T}}. \tag{11.6}$$

Для вычисления (11.6) введем так называемую *сумму состояний*

$$Z = \sum_{N=0}^{\infty} e^{-\varepsilon_N/k_0T}. \tag{11.7}$$

Тогда

$$\bar{\varepsilon} = \frac{\sum \varepsilon_N e^{-\varepsilon_N/k_0T}}{\sum e^{-\varepsilon_N/k_0T}} = -\frac{d}{d\left(\frac{1}{k_0T}\right)} \ln Z, \tag{11.8}$$

что можно непосредственно проверить, дифференцируя выражение (11.7).

С другой стороны, используя (11.5), получим

$$Z = \sum_{N=0}^{\infty} e^{-\varepsilon_N/k_0T} = e^{-\hbar\omega/2k_0T} \sum_{N=0}^{\infty} e^{-N\hbar\omega/k_0T} =$$

$$= e^{-\hbar\omega/2k_0T}\left[1 + e^{-\hbar\omega/k_0T} + e^{-2\hbar\omega/k_0T} + \ldots\right] = \frac{e^{-\hbar\omega/2k_0T}}{1 - e^{-\hbar\omega/k_0T}} \tag{11.9}$$

по формуле для бесконечно убывающей геометрической прогрессии. Подставляя результат (11.9) в выражение (11.8) и дифференцируя по аргументу $1/k_0T$, получим

$$\bar{\varepsilon} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/k_0T} - 1}, \tag{11.10}$$

где член $^1/_2\hbar\omega$, не зависящий от температуры, называется *нулевой энергией осциллятора*. Выражение (11.10) (точнее второе слагаемое в правой части этого выражения) может быть получено, если рассматривать нормальные колебания решетки как квазичастицы—фононы (§ 10, п. 3).

Поскольку операторы рождения и уничтожения фононов удовлетворяют правилам коммутации (10.18), фононы являются бозонами, т. е. описываются в равновесии статистикой Бозе—Эйнштейна.

В некоторых отношениях фононы ведут себя не так, как газ обычных частиц; именно поэтому они называются квазичастицами. Во-первых, при взаимодействии с электронами или друг с другом фононы возникают и исчезают. Во-вторых, среднее число фононов (их концентрация) зависит от температуры. Для газа обычных частиц (атомов, электронов) переменные V, T и N (число частиц) являются независимыми. В случае фононов их число N определяется при заданном V и T из условий равновесия, т. е. из минимума свободной энергии $\mathcal{F}(T, V, N)$.

Таким образом,

$$\left(\frac{\partial \mathcal{F}}{\partial N}\right)_{T,\,V} = \zeta = 0,$$

где ζ по определению—*химический потенциал*.

Равновесное число фононов $\overline{N}$ в одном квантовом состоянии, т. е. в ячейке фазового пространства объема ($1\ см^3 \times h^3$),

с энергией $\hbar\omega$ и химическим потенциалом $\zeta=0$ равно [1])

$$\overline{N}=\frac{1}{e^{\hbar\omega/k_0T}-1}. \tag{11.10a}$$

Очевидно, что их средняя энергия $\overline{\varepsilon}=\hbar\omega\cdot\overline{N}$, что совпадает со вторым слагаемым правой части (11.10).

Отсюда следует, что фононы — это элементарные возбуждения кристалла над его нулевым уровнем энергии $\mathcal{E}_0=\sum\limits_{j,\,q}\frac{\hbar\omega_j(q)}{2}$.

Взаимодействуют с электронами проводимости только эти возбуждения, т. е. нулевые колебания образуют как бы неизменный фон (вакуум) кристалла.

Если, однако, амплитуда нулевых колебаний атомов становится сравнимой с постоянной решетки, то такие *квантовые кристаллы* приобретают ряд интересных свойств; такими свойствами, в определенных условиях, обладают кристаллы твердого гелия.

Используя выражение (11.10), получим в термодинамическом равновесии для полной внутренней энергии кристаллической решетки

$$\mathcal{E}=\mathcal{E}_0+\sum_{j=1}^{3}\sum_{q}\frac{\hbar\omega_{qj}}{e^{\hbar\omega_{qj}/k_0T}-1}+\sum_{j=4}^{3s}\sum_{q}\frac{\hbar\omega_{qj}}{e^{\hbar\omega_{qj}/k_0T}-1}, \tag{11.11}$$

где $\mathcal{E}_0=\sum\limits_{qj}\frac{\hbar\omega_{qj}}{2}$ — не зависящая от T нулевая энергия и выделено суммирование $\sum\limits_{j=1}^{3}$ по трем акустическим ветвям.

3. Будем считать (в грубом приближении), что частоты оптических ветвей не зависят от q и равны своим предельным значениям $\omega_j^0\neq 0$ [2]). В этом случае суммирование по q для каждой оптической ветви эквивалентно умножению на N. Энергию акустических колебаний (второе слагаемое в (11.11)) вычислим в приближении изотропного континуума (§ 9); это приближение тем лучше, чем больше длина волны по сравнению с постоянной решетки.

Используя функцию распределения частот $g(\omega)$ (9.15), представим энергию акустических колебаний в виде

$$\mathcal{E}_{\text{ак}}=\int_0^{\omega_m}\frac{\hbar\omega}{e^{\hbar\omega/k_0T}-1}g(\omega)\,d\omega=\frac{3V\hbar}{2\pi^2v_0^3}\int_0^{\omega_m}\frac{\omega^3\,d\omega}{e^{\hbar\omega/k_0T}-1}, \tag{11.12}$$

[1]) Ансельм А. И., гл. IX, формула (2.9).

[2]) Например, пользуясь (3.6), можно показать, что ширина оптической ветви, т. е. $\omega_{\text{оп}}(0)-\omega_{\text{оп}}\left(\frac{\pi}{a}\right)$ мала, когда масса одного атома много больше массы другого.

где v_0^3 — средняя от обратного куба продольной и поперечной скорости акустических волн (9.16).

Максимальная частота ω_m определяется из условия, что полное число колебаний равняется полному числу нормальных колебаний $3N$ во всех трех акустических ветвях. Таким образом,

$$\int_0^{\omega_m} g(\omega)\, d\omega = \frac{3V}{2\pi^2 v_0^3} \int_0^{\omega_m} \omega^2\, d\omega = \frac{V\omega_m^3}{2\pi^2 v_0^3} = 3N,$$

откуда

$$\omega_m = v_0\left(\frac{6\pi^2}{\Omega_0}\right)^{1/3} \quad \text{и} \quad q_m = \frac{\omega_m}{v_0} = \left(\frac{6\pi^2}{\Omega_0}\right)^{1/3}. \tag{11.13}$$

Здесь $\Omega_0 = V/N$ — объем элементарной ячейки, а q_m — максимальное значение волнового вектора. Определим «постоянную решетки» равенством $\Omega_0 = a^3$. По порядку величины $\omega_m \sim v_0/a$ и, следовательно, максимальное значение волнового вектора и минимальное значение длины волны: $q_m = \omega_m/v_0 \sim 1/a$ и $\lambda_{\min} = 2\pi/q_m \sim a$. Для кубического кристалла бриллюэновская зона имеет форму куба с ребром, равным $2\pi/a$, так что максимальные значения прямоугольных составляющих вектора $\boldsymbol{q}$ равны π/a (5.9а).

В приближении упругого континуума теории Дебая возможная область значений $\boldsymbol{q}$ заключена в сферу радиуса q_m (11.13).

Определим характеристическую температуру твердого тела (*температуру Дебая*) равенством

$$T_c = \frac{\hbar\omega_m}{k_0} = \left(\frac{6\pi^2}{\Omega_0}\right)^{1/3} \frac{\hbar}{k_0} v_0. \tag{11.14}$$

Так как $\omega_m \sim \frac{v_0}{a} \sim \frac{10^5}{10^{-8}} \sim 10^{13}$ и $k_0 \sim 10^{-16}$, то $T_c \sim 100\,^\circ\text{K}$. Можно ввести температуры Дебая, соответствующие предельным частотам оптических ветвей:

$$T_{cj} = \hbar\omega_j^0/k_0, \tag{11.14а}$$

которые тоже порядка $10^2 - 10^3$ градусов Кельвина, но, вообще говоря, $T_{cj} > T_c$.

Вводя в выражение (11.12) переменную интегрирования $x = \hbar\omega/k_0T$ и используя определения характеристических температур (11.14) и (11.14а), имеем

$$\mathcal{E} = \mathcal{E}_0 + Nk_0T\left\{3D\left(\frac{T_c}{T}\right) + \sum_{i=4}^{3s} \frac{T_{cj}/T}{e^{T_{cj}/T} - 1}\right\}, \tag{11.15}$$

где *функция Дебая*

$$D(t) = \frac{3}{t^3} \int_0^t \frac{x^3\, dx}{e^x - 1}. \tag{11.15а}$$

Рассмотрим вначале высокие температуры, когда $T \gg T_{cj}$ и тем более $T \gg T_c$, так как $T_c < T_{cj}$. В этом случае аргумент функции Дебая $t = T_c/T \ll 1$. Заменяя в подынтегральной функции в выражении (11.15а) $e^x \approx 1 + x$, легко видеть, что $D(t) \approx 1$. Разлагая также экспоненту в членах, соответствующих оптическим ветвям, получим

$$\sum_{j=4}^{3s} \frac{T_{cj}/T}{e^{T_{cj}/T} - 1} \approx \sum_{j=4}^{3s} 1 = 3s - 3.$$

Таким образом, $\mathcal{E} = \mathcal{E}_0 + 3sNk_0T$, а теплоемкость $c_V = \partial\mathcal{E}/\partial T = 3sNk_0$ в соответствии с законом Дюлонга и Пти (11.4).

Рассмотрим теперь случай низких температур, когда $T \ll T_c$ и тем более $T \ll T_{cj}$. Пренебрегая величинами порядка $e^{-T_c/T}$ по сравнению с единицей, отбросим в выражении (11.15) члены, соответствующие оптическим ветвям, и заменим в интеграле функции $D(t)$ верхний предел на ∞. Так как[1])

$$\int_0^\infty \frac{x^3\,dx}{e^x - 1} = \frac{\pi^4}{15},$$

то

$$\mathcal{E} = \mathcal{E}_0 + \frac{3\pi^4 N k_0 T^4}{5T_c^3} = \mathcal{E}_0 + \frac{\pi^2 V (k_0 T)^4}{10\hbar^3 v_0^3}, \qquad (11.16)$$

т. е. зависящая от температуры часть энергии пропорциональна T^4. Теплоемкость

$$c_V = \frac{\partial\mathcal{E}}{\partial T} = \frac{12\pi^4 k_0}{5}\left(\frac{T}{T_c}\right)^3, \qquad (11.17)$$

т. е. пропорциональна T^3. «Закон T^3» довольно хорошо оправдывается на опыте при температурах порядка 20—50 °К. Теория теплоемкости Дебая, основанная на выражениях для $g(\omega)$ (9.15) и для предельной частоты (11.13), должна хорошо оправдываться для низких температур, когда возбуждены только длинные волны и, следовательно, применима аппроксимация упругого континуума. С другой стороны, при высоких температурах, когда теплоемкость определяется просто числом степеней свободы решетки, теория Дебая приводит к правильному результату — закону Дюлонга и Пти. Для промежуточных температур выражение для теплоемкости c_V, следующее из теории Дебая, может рассматриваться только как более или менее удачная интерполяционная формула. Дифференцируя выражение (11.15) по T, получим

$$c_V = \frac{\partial\mathcal{E}}{\partial T} = 3Nk_0\left\{4D\left(\frac{T_c}{T}\right) - \frac{3\,(T_c/T)}{e^{T_c/T} - 1}\right\} + Nk_0 \sum_{j=4}^{3s} \frac{e^{T_{cj}/T}(T_{cj}/T)^2}{\left(e^{T_{cj}/T} - 1\right)^2}. \qquad (11.18)$$

[1]) Ландау Л. Д., Лифшиц Е. М./Теоретическая физика, т. 5.— Статистическая физика.—3 изд.—М., 1976, с. 223.

Температурная зависимость теплоемкости акустических ветвей характеризуется величиной

$$\frac{c_V^{(D)}}{3k_0N}=\left\{4D\left(\frac{T_c}{T}\right)-\frac{3(T_c/T)}{e^{T_c/T}-1}\right\}=3\left(\frac{T}{T_c}\right)^3\int_0^{T_c/T}\frac{x^4\,dx}{(e^x-1)^2}\,dx, \quad (11.18a)$$

где правая часть может быть получена интегрированием по частям интеграла, определяющего функцию $D(T_c/T)$. На рис. III.16 построена величина (11.18а) как функция T/T_c. При низких температурах она пропорциональна $(T/T_c)^3$, при высоких — стремится к единице. Отклонение выражения (11.18а) от единицы определяет отступление теплоемкости акустических ветвей от классического значения $3Nk_0$. Из кривой видно, что практически характеристической температурой, отделяющей классическую область от квантовой, является не T_c, а скорее $T_c/3$. Таким образом, при температурах $T\geqslant T_c$ справедливы классические выражения для теплоемкости, а при $T<T_c/10$ должны действовать квантовые законы.

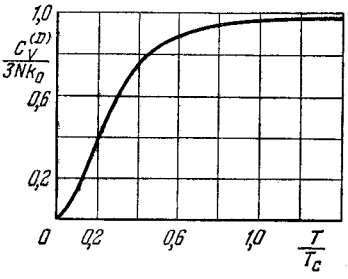


Рис. III.16.

Теплоемкости, соответствующие оптическим ветвям, часто называют *эйнштейновскими членами*[1]). При $T\to 0$ они убывают не по закону T^3, а быстрее, как $T^{-2}e^{-T_{cj}/T}$. В том случае, когда $T_{cj}\gg T_c$ (молекулярные решетки), возможны случаи, когда $T\geqslant T_c$ и $T\ll T_{cj}$, так что акустические ветви полностью возбуждены, а теплоемкостью оптических ветвей можно пренебречь. В этом случае решетка ведет себя как классическая одноатомная с массами «атомов» $M=\sum_{k=1}^{s}m_k$.

4. Более глубокое экспериментальное и теоретическое изучение теплоемкости твердых тел при низких температурах показало, что истинная область закона T^3 простирается только на несколько градусов вблизи абсолютного нуля. Видимость же выполнения закона T^3 при температурах 20—50 °К обусловлена другими причинами, рассмотренными ниже. Конкретные расчеты показали, что даже при низких температурах $\left(T<\frac{T_c}{10}\right)$ теп-

[1]) Теплоемкость одноатомного твердого тела была впервые с квантовой точки зрения рассмотрена А. Эйнштейном; при этом им было сделано упрощающее предположение, что все атомы решетки колеблются с одной определенной частотой. Очевидно, что при этом он получил выражение для теплоемкости, совпадающее с теплоемкостью одной оптической ветви в выражении (11.18) (умноженной на 3).

лоемкость остается чувствительной к дискретности структуры решетки.

Представим себе, что нами экспериментально или теоретически определена точная зависимость теплоемкости c_V от температуры T. Используя теорию теплоемкости Дебая, основанную на представлении об упругом континууме, положим

$$c_V(T) = c_V^{(D)}(T_c/T), \tag{11.19}$$

где $c_V^{(D)}$ — дебаевская теплоемкость (11.18а), зависящая от температуры и одного параметра T_c [1]). Отступление $c_V(T)$ от теории теплоемкости Дебая можно формально описать, считая характеристическую температуру T_c зависящей от температуры T.

Так, например, Келлерман [2]) подробно рассмотрел колебания решетки NaCl, вычисляя коэффициенты $D_{\alpha\beta}^{kk'}(\boldsymbol{q})$ по формуле (5.10а) и решая характеристическое уравнение (5.11) для частоты ω^2. При определении кулоновских сил взаимодействия ионы рассматривались как точечные заряды. Силы отталкивания между ионами определялись из данных по сжимаемости.

В общем случае число колебаний во всех $3s$ ветвях на малый интервал частоты $\Delta\omega$ равно [3])

$$g(\omega)\,\Delta\omega = \frac{V}{(2\pi)^3}\sum_{j=1}^{3s}\iiint\limits_{\omega \leqslant \omega_j(q_x,\, q_y,\, q_z) \leqslant \omega+\Delta\omega} dq_x\, dq_y\, dq_z. \tag{11.20}$$

Для определения отсюда функции распределения колебаний $g(\omega)$ необходимо знать для всех ветвей зависимость $\omega_j(q_x, q_y, q_z)$, что может быть сделано только путем численных расчетов. Выражение (11.20) может быть использовано вместо приближенного выражения (9.15) для определения внутренней энергии кристалла (11.12).

На рис. III.17 и III.18 представлены, по Келлерману, функция распределения колебаний $g(\omega)$ для акустических ветвей и следующая из нее и из теории Дебая зависимость T_c от температуры [4]). Мы видим, что $g(\omega)$ весьма существенно отличается от параболического закона (9.15), представленного на рис. III.17 пунктирной кривой. Из дебаевской формулы (11.17) видно, что закон T^3 реализуется только тогда, когда T_c достигает при низких температурах постоянного значения, т. е., как видно из рис. III.18, ниже 5—7 °K. С другой стороны, в интервале

---

[1]) Мы рассматриваем только акустические ветви, считая что оптические ветви не возбуждены в интересующей нас области температур.

[2]) Kellermann E. W.—Phil. Trans. Roy. Soc.—1940 v. 238, p. 513;—Proc. Roy. Soc.—1941. v. A 178, p. 17.

[3]) При этом надо иметь в виду, что, вообще говоря, не все колебательные ветви обязательно вносят что-либо в данном интервале ω, $\omega+\Delta\omega$.

[4]) Кружки на рис III.18—экспериментальные данные.

20—50°К у T_c наблюдается пологий минимум, вблизи которого T_c можно считать постоянной и, следовательно, ожидать приближенного выполнения закона T^3.

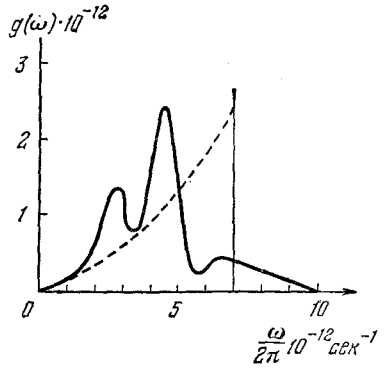


Рис. III.17.

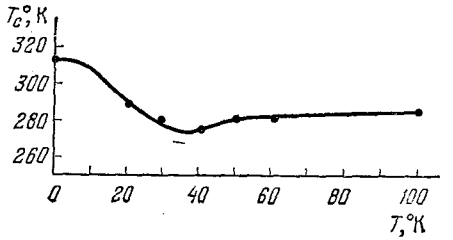


Рис. III.18.

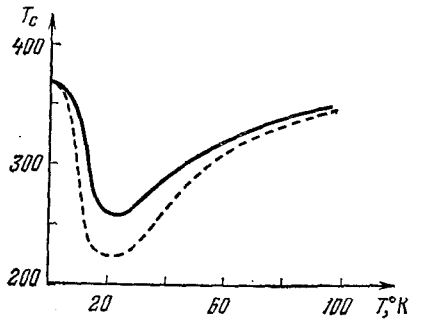


Рис. III.19.

Келлерман использовал для расчета теплоемкости кристалла NaCl эмпирические данные по его сжимаемости. Более прямой путь — непосредственное использование функции распределения частот $g(\omega)$, полученной из опытов по рассеянию нейтронов в твердом теле.

Ф. Джонсон и В. Кочран[1]) определили функцию $g(\omega)$ для германия из кривых дисперсии $\omega_j = \omega_j(\boldsymbol{q})$, полученных на основании опытов по рассеянию нейтронов. Вычисленная ими для германия зависимость температуры Дебая T_c от температуры T (пунктирная кривая на рис. III.19) может быть сравнена с этой же зависимостью, полученной из теплоемкости (сплошная кривая). Совпадение, как мы видим неполное, но вполне удовлетворительное. Расхождение мы относим за счет неточного определения функции распределения частот $g(\omega)$.

§ 12. Уравнение состояния твердого тела

1. Выведем уравнение состояния, т. е. установим связь между давлением P, объемом V и температурой T твердого тела. Для простоты будем рассматривать одноатомное твердое тело в дебаевском приближении.

[1]) Johnson F. A., Cochran W.—Proc. Exeter Conf. on Semiconductors, 1962, p. 498.

Наиболее прямой путь получения уравнения состояния заключается в вычислении свободной энергии системы[1])

$$\mathcal{F} = -k_0 T \ln Z, \tag{12.1}$$

где сумма состояний

$$Z = \sum_n e^{-\varepsilon_n/k_0 T} \Omega(\varepsilon_n). \tag{12.1a}$$

Здесь $\Omega(\varepsilon_n)$ — статистический вес или число различных состояний системы с одной и той же энергией ε_n; другими словами, $\Omega(\varepsilon_n)$ — кратность вырождения энергетического уровня ε_n. Для линейного гармонического осциллятора $\Omega(\varepsilon_n) = 1$ и Z равно выражению (11.9). При непрерывном изменении энергии ε сумма состояний заменяется интегралом состояний.

Из термодинамики известно[2]), что давление

$$P = -\left(\frac{\partial \mathcal{F}}{\partial V}\right)_T, \tag{12.2}$$

где производная по V берется при постоянном T.

Свободная энергия осциллятора при частоте ω равна

$$\mathcal{F}_{\text{осц}} = \frac{\hbar\omega}{2} + k_0 T \ln\left(1 - e^{-\hbar\omega/k_0 T}\right), \tag{12.3}$$

как это следует из выражений (12.1) и (11.9).

Для твердого тела, описываемого $3G^3 = 3N$ независимыми нормальными колебаниями (осцилляторами), свободная энергия равна сумме свободных энергий осцилляторов:

$$\mathcal{F} = \sum \mathcal{F}_{\text{осц}} = \sum_{qj} \frac{\hbar\omega_{qj}}{2} + k_0 T \sum_{qj} \ln\left(1 - e^{-\hbar\omega_{qj}/k_0 T}\right). \tag{12.4}$$

В приближении дебаевского континуума

$$\mathcal{F} = \mathcal{E}_0 + k_0 T \frac{3V}{2\pi^2 v_0^3} \int_0^{\omega_m} \ln\left(1 - e^{-\hbar\omega/k_0 T}\right) \omega^2 \, d\omega =$$

$$= \mathcal{E}_0 + 9Nk_0 T \left(\frac{T}{T_c}\right)^3 \int_0^{T_c/T} \ln\left(1 - e^{-x}\right) x^2 \, dx, \tag{12.4a}$$

если воспользоваться соотношениями (9.15) и (11.14).

При определении давления по формуле (12.2) необходимо (12.4а) продифференцировать по объему V при постоянной температуре T. Можно показать, что характеристическая температу-

[1]) Ансельм А. И. Основы статистической физики и термодинамики.—М., 1973, с. 53, 111.

[2]) См. там же.

ра T_c или пропорциональная ей предельная частота ω_m в ангармоническом приближении зависят от объема V. Проиллюстрируем это на примере простой одномерной решетки, предельная частота колебаний которой, согласно (2.5а), равна $\omega_m^2 = \frac{4}{m}\beta$, где β — коэффициент квазиупругой силы.

Согласно (1.3) сила взаимодействия атомов в ангармоническом приближении равна

$$F = -\beta x + \gamma x^2.$$

Здесь $x = R - R_0$ (R_0 — равновесное расстояние между атомами), $\beta = \mathcal{U}''(R_0) > 0$ и $\gamma = -{}^1/_2\mathcal{U}'''(R_0) > 0$; $\mathcal{U}(R)$ — потенциальная энергия взаимодействия атомов.

При однородном растяжении или сжатии атомной цепочки на каждый атом дополнительно действует некоторое однородное поле $A = \text{const}$, энергия атома в котором равна AR. Так как сила в новом положении равновесия $R_0 + \Delta R_0$ равна нулю, то $\mathcal{U}'(R_0 + \Delta R_0) + A = 0$. Разлагая $\mathcal{U}'(R_0 + \Delta R_0)$ в ряд по степеням ΔR_0 и учитывая, что $\mathcal{U}'(R_0) = 0$, получим $\mathcal{U}''(R_0)\Delta R_0 + A = 0$ или $A = -\beta\Delta R_0$. С другой стороны, измененная предельная частота $\omega_m + \Delta\omega_m$ удовлетворяет соотношению

$$(\omega_m + \Delta\omega_m)^2 = \frac{4}{m}\mathcal{U}''(R_0 + \Delta R_0)$$

(вторая производная по R от энергии AR равна нулю). Разлагая правую часть по ΔR_0, пренебрегая слева $(\Delta\omega_m)^2$ и используя значение ω_m^2, получим

$$\frac{\Delta\omega_m}{\omega_m} = -\frac{\gamma\Delta R_0}{\beta}.$$

Мы видим, что предельная частота ω_m, а следовательно, и T_c меняются при изменении «объема» ΔR_0 только в ангармоническом приближении ($\gamma \neq 0$). Так как β и γ положительны, то при увеличении «объема» ΔR_0 предельная частота уменьшается ($\Delta\omega_m < 0$) и наоборот.

Таким образом, давление

$$P = -\left(\frac{\partial\mathcal{F}}{\partial V}\right)_T = -\frac{\partial\mathcal{E}_0}{\partial V} - 3Nk_0TD\left(\frac{T_c}{T}\right)\frac{1}{T_c}\frac{\partial T_c}{\partial V}, \qquad (12.5)$$

если преобразовать интеграл в выражении (12.4а) по частям и воспользоваться обозначением (11.15а).

Введем в рассмотрение *параметр Грюнейзена*

$$\gamma_G \equiv -\frac{V}{T_c}\frac{dT_c}{dV} = -\frac{d\omega_m/\omega_m}{dV/V} = -\frac{d\ln\omega_m}{d\ln V} > 0, \qquad (12.6)$$

не зависящий от температуры. Из определения (12.6) видно, что он непосредственно связан с ангармонизмом взаимодействия

атомов. В простом одномерном случае, рассмотренном выше,

$$\gamma_G = -\frac{\Delta\omega_m/\omega_m}{\Delta R_0/R_0} = \frac{\gamma R_0}{\beta}. \tag{12.6а}$$

В гармоническом приближении ($\gamma = 0$) параметр Грюнейзена $\gamma_G = 0$. Обозначая зависящую от температуры часть внутренней энергии через $\mathscr{E}_T = 3Nk_0TD(T_c/T)$, получим *уравнение состояния твердого тела*

$$P = -\frac{\partial \mathscr{E}_0}{\partial V} + \frac{\gamma_G \mathscr{E}_T}{V}, \tag{12.7}$$

где первый член правой части от T не зависит.

2. Выведем из уравнения состояния так называемое *соотношение Грюнейзена*. Дифференцируя выражение (12.7) по T и учитывая, что $\left(\frac{\partial \mathscr{E}_T}{\partial T}\right)_V = c_V$, получим

$$\left(\frac{\partial P}{\partial T}\right)_V = \frac{\gamma_G c_V}{V}. \tag{12.8}$$

Воспользуемся термодинамическим тождеством[1])

$$\left(\frac{\partial P}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_P \left(\frac{\partial T}{\partial P}\right)_V = -1 \tag{12.9}$$

и запишем его в виде

$$\left(\frac{\partial V}{\partial T}\right)_P = -\left(\frac{\partial V}{\partial P}\right)_T \left(\frac{\partial P}{\partial T}\right)_V, \tag{12.9а}$$

что возможно, так как $\left(\frac{\partial P}{\partial V}\right)_T = 1 : \left(\frac{\partial V}{\partial P}\right)_T$ и т. д.

Вводя коэффициент линейного теплового расширения

$$\alpha = \frac{1}{3V}\left(\frac{\partial V}{\partial T}\right)_P \tag{12.10}$$

и изотермическую сжимаемость

$$k = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T, \tag{12.11}$$

получим из равенств (12.9а), (12.10), (12.11) и (12.8) соотношение Грюнейзена

$$3V\alpha = \gamma_G k c_V. \tag{12.12}$$

Опыт показывает, что, действительно, левая и правая части этого равенства имеют одинаковую температурную зависимость, если считать, что γ_G от температуры не зависит. Изучая на опыте изменение сжимаемости k при больших давлениях P, можно независимо определить постоянную Грюнейзена γ_G, которая хорошо совпадает с вычисленной из уравнения (12.12).

[1]) Ансельм А. И. Основы статистической физики и термодинамики.— М., 1973, с. 117.

3. Рассмотрим газ равновесных фононов в континуальном (дебаевском) приближении. Число равновесных фононов в интервале частот $(\omega, \omega+d\omega)$ равно

$$dN_\omega = \bar{N} g(\varepsilon)\, d\omega = \frac{3V}{2\pi^2 v_0^3} \frac{\omega^2\, d\omega}{e^{\hbar\omega/k_0T}-1}. \qquad (12.13)$$

Здесь $\bar{N}$ — среднее число фононов в элементарной ячейке фазового пространства (11.10а), а $g(\omega)$ — функция распределения частот в континуальном приближении (9.15). Энергия этих фононов

$$d\mathscr{E}_\omega = \hbar\omega\, dN_\omega = \frac{3V\hbar}{2\pi^2 v_0^3} \frac{\omega^3\, d\omega}{e^{\hbar\omega/k_0T}-1}. \qquad (12.14)$$

Если заменить в этом выражении множитель 3 на 2, в соответствии с только двумя возможными (поперечными) поляризациями фотона, и скорость звука v_0 — на скорость света c, то получим известную формулу Планка для распределения энергии в спектре черного излучения. Полное число фононов в объеме V равно

$$N = \int dN_\omega = \frac{3V}{2\pi^2 v_0^3} \int_0^{\omega_m} \frac{\omega^2\, d\omega}{e^{\hbar\omega/k_0T}-1} = \frac{3V}{2\pi^2 v_0^3} \left(\frac{k_0T}{\hbar}\right)^3 \int_0^{T_c/T} \frac{x^2\, dx}{e^x-1}, \qquad (12.15)$$

где T_c — температура Дебая.

Из выражения (12.15) следует, что при $T \ll T_c$, когда с точностью до величин порядка $e^{-T_c/T}$ можно заменить верхний предел интеграла на ∞, число фононов $N \sim VT^3$. Для высоких температур, когда $T \gg T_c$ и можно под знаком интеграла положить $e^x - 1 \approx x$, число фононов $N \sim VT$. На рис. III.20 представлена зависимость концентрации фононов N/V от температуры T.

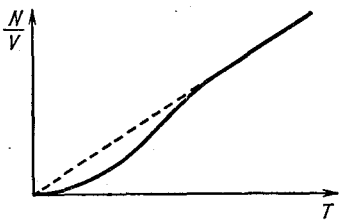


Рис. III. 20.

Полная энергия фононов в объеме V равна

$$\mathscr{E} = \int d\mathscr{E}_\omega = \frac{3V\hbar}{2\pi^2 v_0^3} \int_0^{\omega_m} \frac{\omega^3\, d\omega}{e^{\hbar\omega/k_0T}-1}, \qquad (12.16)$$

что в точности совпадает с энергией нормальных колебаний (11.12). Таким образом, энергия нормальных колебаний решетки совпадает с энергией фононов, распределенных по закону Бозе—Эйнштейна.

В термодинамике показывается [1]), что свободная энергия

$$\mathscr{F} = -T \int \frac{\mathscr{E}(T)\, dT}{T^2}, \qquad (12.17)$$

[1]) А. И. Ансельм, с. 110.

где $\mathcal{E}(T)$ — внутренняя энергия системы. Неопределенный интеграл в (12.17) соответствует неопределенной аддитивной константе в выражении для свободной энергии. Если подставить в (12.17) $\mathcal{E}(T)$ из уравнения (12.16) и произвести интегрирование по T, то получим свободную энергию, совпадающую с $\mathcal{F}-\mathcal{E}_0 \equiv \mathcal{F}_T$ в выражении (12.4а). Мы можем теперь смотреть на $\mathcal{F}_T$ в (12.4а) как на свободную энергию фононного газа, а на уравнения (12.5) и (12.7) без члена $\frac{\partial \mathcal{E}_0}{\partial V}$ — как на уравнения состояния фононного газа. Как мы указывали выше, в гармоническом приближении параметр Грюнейзена $\gamma_G = 0$, поэтому в этом приближении давление фононного газа равно нулю.

В ангармоническом приближении давление фононного газа

$$P_{\text{фон}} = \gamma_G \frac{\mathcal{E}_T}{V}, \tag{12.18}$$

как это следует из выражения (12.7).

Анализ вопроса показывает, что поток импульса, связанный с плоской гармонической бегущей волной с определенным квазиимпульсом $\hbar \boldsymbol{q}$, равен нулю, т. е. фонон не обладает импульсом (количеством движения). В этом фонон существенно отличается от фотона, который обладает импульсом $\hbar \boldsymbol{k}$ ($\boldsymbol{k}$ — волновой вектор фотона).

§ 13. Тепловое расширение и теплопроводность твердого тела

Мы объединили в одном параграфе явления теплового расширения и теплопроводности твердого тела, так как они оба определяются ангармонической частью сил взаимодействия атомов. Как тепловое расширение, так и тепловое сопротивление[1]) (равное $1/\varkappa$, где $\varkappa$ — коэффициент теплопроводности) исчезают, если положить коэффициент ангармоничности $\gamma = 0$.

Мы рассмотрим тепловое расширение на простой модели двух взаимодействующих атомов. Эта модель дает возможность выяснить не только принципиальную сторону явления, но и позволяет определить правильный порядок величины коэффициента теплового расширения.

В теплопроводности мы ограничимся некоторыми общими соображениями и определим коэффициент теплопроводности при высоких температурах из соображений размерности.

1. Рассмотрим два атома, которые при малых отклонениях от положения равновесия $x = R - R_0$ взаимодействуют друг с другом по закону (1.3), т. е. с силой

$$F = -\frac{dU}{dx} = -\beta x + \gamma x^2 \tag{13.1}$$

[1]) Обусловленные фонон-фононным взаимодействием.

и потенциальной энергией

$$U=\frac{1}{2}\beta x^2-\frac{1}{3}\gamma x^3. \qquad (13.1a)$$

Вероятность отклонения атома от положения равновесия на величину x равна по Больцману

$$f(x)=A\exp\left(-\frac{U}{k_0T}\right)\approx Ae^{-\beta x^2/2k_0T}\left(1+\frac{\gamma x^3}{3k_0T}\right), \qquad (13.2)$$

где экспонента, соответствующая ангармоническому члену, разложена в ряд

$$\exp\left(\frac{\gamma x^3}{3k_0T}\right)\approx 1+\frac{\gamma x^3}{3k_0T}.$$

Постоянная A в выражении (13.2) определяется из условия нормировки:

$$\int_{-\infty}^{+\infty} f(x)\,dx=A\int_{-\infty}^{+\infty} e^{-\beta x^2/2k_0T}\left(1+\frac{\gamma x^2}{3k_0T}\right)dx=1.$$

Интеграл от второго слагаемого, пропорционального γ, равен нулю ввиду нечетности подынтегральной функции; поэтому (Приложение 7)

$$A=\left(\frac{\beta}{2\pi k_0T}\right)^{1/2}.$$

Среднее отклонение атома от положения равновесия

$$\bar{x}=\int_{-\infty}^{+\infty} xf(x)\,dx=$$

$$=\left(\frac{\beta}{2\pi k_0T}\right)^{1/2}\int_{-\infty}^{+\infty} e^{-\beta x^2/2k_0T}\left(x+\frac{\gamma x^4}{3k_0}\right)dx=\frac{\gamma k_0T}{\beta^2}, \qquad (13.3)$$

где интеграл от первого слагаемого, содержащего множитель x, вновь равен нулю ввиду нечетности подынтегральной функции. Интеграл от второго слагаемого вычисляется элементарно (Приложение 7).

По определению коэффициент линейного теплового расширения α есть удлинение в расчете на единицу длины и на 1°C; таким образом,

$$\alpha=\frac{\bar{x}}{aT}=\frac{\gamma k_0}{a\beta^2}, \qquad (13.4)$$

где $a=R_0$ — постоянная решетки. Мы видим, что коэффициент теплового расширения пропорционален коэффициенту ангармоничности и при $\gamma=0$ равен нулю.

Для примера рассмотрим одновалентный ионный кристалл. В этом случае можно положить:

$$F=-\frac{e^2}{R^2}+\frac{B}{R^{10}}. \qquad (13.5)$$

Здесь $-e^2/R^2$ — кулоновское притяжение между соседними разноименными недеформируемыми ионами, B/R^{10} — сила отталкивания между этими же ионами, быстро возрастающая при уменьшении расстояния между ними R (пропорциональность этой силы R^{-10} достаточно хорошо аппроксимирует экспоненциальную зависимость, получаемую при квантовомеханическом расчете). В равновесии $F=0=-\frac{e^2}{a^2}+\frac{B}{a^{10}}$, где $a=R_0$ — равновесное расстояние между ближайшими ионами. Отсюда $B=e^2a^8$.

Так как $R=a+x$, то для малых x

$$F=-\frac{e^2}{(a+x)^2}+\frac{e^2a^8}{(a+x)^{10}}\approx-\frac{8e^2}{a^3}x+\frac{52e^2}{a^4}x^2. \qquad (13.6)$$

Из сравнения с уравнением (13.1) следует:

$$\beta=8e^2/a^3, \quad \gamma=52e^2/a^4. \qquad (13.6а)$$

Подставляя результат (13.6а) в выражение (13.4), получим

$$\alpha=52ak_0/64e^2. \qquad (13.6б)$$

Для $a=3\cdot10^{-8}$ *см*, $k_0=1{,}38\cdot10^{-16}$ *эрг/град*, $e=4{,}8\cdot10^{-10}$ *CGSE*, получим $\alpha=1{,}5\cdot10^{-5}$ *град*$^{-1}$, что дает правильную по порядку величину.

2. Впервые Дебай (1914) показал, что тепловое сопротивление в твердом теле обусловлено ангармонизмом колебаний атомов и что, если учитывать колебания только в гармоническом приближении, то тепловое сопротивление равно нулю. Это утверждение представляется достаточно наглядным — для гармонических волн имеет место принцип линейной суперпозиции, согласно которому волны распространяются в кристалле независимо, не рассеиваясь друг на друге. В такой модели тепловое сопротивление равно нулю, так как тепловой поток распространяется со скоростью звука. Ввиду того, что плоской гармонической волне с определенным волновым вектором $\boldsymbol{q}$ соответствует фонон с квазиимпульсом $\hbar\boldsymbol{q}$ и энергией $\hbar\omega_q$, можно сказать, что в гармоническом приближении фононы не взаимодействуют, т. е. не сталкиваются друг с другом. В общем случае в кристаллической решетке ангармонизм учитывается членами третьей степени в смещениях атомов $u^k_{n\alpha}$ в разложении потенциальной энергии $\Phi(u)$ (см. (5.2)). Теория показывает, что если учитывать анагармонические члены в потенциальной энергии $\Phi(u)$ как малое возмущение, то это приводит к представлению о возможности одновременного «столкновения» трех фононов. При этом процессы столкновения имеют следующий характер — либо два фонона превращаются в один, либо один фонон распадается на два. Таким образом, в процессе «столкновения» квазичастицы фононы рождаются и исчезают. Последовательная теория тепло-

проводности кристаллов, основанная на кинетическом уравнении для фононов, была развита Пайерлсом (1929).

В этом параграфе мы ограничимся определением коэффициента теплопроводности $\varkappa$ для высоких температур $T \gg T_c$ из соображений размерности. Как теория Дебая, так и другие более последовательные теории приводят к заключению, что при высоких температурах ($T \gg T_c$) коэффициент теплопроводности $\varkappa$ обратно пропорционален абсолютной температуре T. С другой стороны, коэффициент теплопроводности пропорционален длине свободного пробега фонона l, а последняя обратно пропорциональна вероятности рассеяния фонона, которая в свою очередь пропорциональна квадрату матричного элемента энергии возмущения. Так как энергия ангармонического возмущения пропорциональна γ и, следовательно, квадрат матричного элемента пропорционален γ^2, то коэффициент теплопроводности обратно пропорционален γ^2. Таким образом,

$$\varkappa \sim \frac{1}{\gamma^2 T}. \tag{13.7}$$

При высоких температурах, когда несущественны квантовые эффекты, теплопроводность $\varkappa$ для простой модели твердого тела, по Дебаю, может еще зависеть от следующих величин: постоянной Больцмана k_0, постоянной решетки a, массы атома кристалла M и скорости звука v_0.

Так как $v_0 = \text{const} \cdot a\sqrt{\beta/M}$, где β — коэффициент квазиупругой силы, то β не может фигурировать одновременно в качестве независимого параметра.

Таким образом, задача сводится к определению показателей степеней в размерном уравнении

$$\varkappa = \text{const} \cdot \frac{k_0^l a^m M^n v_0^p}{\gamma^2 T}, \tag{13.8}$$

где const — безразмерная постоянная.

Применяя П-теорему Бриджмэна[1]) (или просто подбирая) можно показать, что для одинаковой размерности левой и правой частей в выражении (13.8) требуется, чтобы

$$l = 0, \quad m = -8, \quad n = +3, \quad p = +7. \tag{13.8а}$$

Таким образом,

$$\varkappa = \text{const} \cdot \frac{M^3 v_0^7}{a^8 \gamma^2 T}. \tag{13.9}$$

Учитывая выражение (13.4), а также связь между величинами β и v_0, a и M, можно исключить из формулы (13.9) неизвестный

[1]) Бриджмэн П. В. Анализ размерностей.—М.—Л.: ГТТИ, 1934.

коэффициент γ^2:

$$\varkappa = \text{const} \cdot \frac{k_0^2}{a^2 M v_0 a^2 T}, \qquad (13.9a)$$

где безразмерная постоянная const иная, чем в выражении (13.9).

Формула (13.9а) может быть сравнена с опытом после того, как определена const для какого-либо одного вещества. Вместо скорости звука v_0 можно в выражениях (13.9) и (13.9а) ввести температуру Дебая T_c по формуле (11.14).

Необходимо учитывать приближенный характер выражений (13.9) и (13.9а), так как потенциальная энергия кристалла определяется здесь только двумя силовыми константами β и γ (или v_0 и γ). Однако в рамках такой модели формулы (13.9) и (13.9а) абсолютно точны, и было бы бесполезно пытаться с тем же набором параметров, определяющих свойства кристалла, получить более точное выражение для коэффициента теплопроводности.

ГЛАВА IV

ЭЛЕКТРОНЫ В ИДЕАЛЬНОМ КРИСТАЛЛЕ

§ 1. Общая постановка задачи. Адиабатическое приближение

1. Любое твердое тело макроскопических размеров представляет собой совокупность огромного числа атомных ядер и электронов. Обычно атомные ядра одного элемента представляют собой естественную смесь его изотопов, однако для большинства свойств твердого тела это обстоятельство не является существенным. Важной особенностью твердого тела, установленной эмпирически, является то, что атомные ядра занимают в нем более или менее фиксированные положения в пространстве. В идеальном кристалле эти положения образуют трехмерную периодическую решетку. В нейтральном твердом теле общий положительный заряд ядер равен абсолютной величине суммарного заряда всех электронов.

В некоторых случаях, особенно при изучении границы (поверхности) кристалла, вблизи которой потенциальная энергия электронов проводимости имеет другое значение, чем в объеме, приходится рассматривать заряженные области тела, в которых объемный заряд и электростатический потенциал самосогласованы друг с другом.

Если не интересоваться такими процессами в твердом теле, при которых имеют место ядерные превращения и становятся существенными релятивистские эффекты и спин электрона и ядер [1]), то стационарные состояния системы описываются уравнением Шредингера

$$\hat{\mathcal{H}}\Psi = W\Psi \tag{1.1}$$

с гамильтонианом

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m}\sum_i \nabla_{r_i}^2 - \frac{\hbar^2}{2}\sum_J \frac{1}{M_J}\nabla_{R_J}^2 + V(r, R), \tag{1.1a}$$

[1]) Известно, что ряд явлений в кристаллах (спин-орбитальное расщепление электронного спектра, парамагнитный и ядерный резонанс и др.) существенно зависит от спина электронов и ядер.

где кулоновская энергия взаимодействия электронов и ядер

$$V(\boldsymbol{r}, R)=\sum_{J<K}\frac{Z_J Z_K e^2}{R_{JK}}+\sum_{i<k}\frac{e^2}{r_{ik}}-\sum_{iJ}\frac{Z_J e^2}{r_{iJ}}. \qquad (1.1б)$$

Здесь m — масса электрона, M_J — масса J-го ядра, $\boldsymbol{r}_i$ и $\boldsymbol{R}_J$ — радиусы-векторы i-го электрона и J-го ядра, R_{JK}, r_{ik} и r_{iJ} — расстояния между соответствующими ядрами и электронами, Z_J — атомный номер J-го ядра.

Спектр собственных значений энергии системы W имеет, вообще говоря, сложный характер.

Так как поведение физической системы полностью определяется ее волновой функцией, то решение уравнения (1.1) позволило бы в принципе ответить на все вопросы, связанные со свойствами твердого тела, в частности на вопросы: почему совокупность данных ядер и электронов образует кристаллическую решетку того или другого типа, каковы тепловые, электрические, магнитные и оптические свойства данного твердого тела и т. д.

Так как макроскопический образец твердого тела содержит около 10^{23} частиц и, следовательно, волновая функция уравнения (1.1) зависит от такого же числа переменных, то, конечно, практически (даже написать!) решение уравнения (1.1) невозможно. Более того, если бы мы даже нашли способ написания (задания) волновой функции от такого числа переменных, то это было бы бесполезно, так как использование ее для вычисления наблюдаемых на опыте величин натолкнулось бы на затруднения, которые следовало бы признать принципиальными.

Задача физической теории как раз и заключается в отыскании таких достаточно обоснованных приближений, которые позволили бы интерпретировать и вычислить величины, наблюдаемые на опыте.

2. Попробуем упростить уравнение (1.1), воспользовавшись тем, что масса электрона m много меньше массы атомных ядер M_J (адиабатическое приближение). Рассматривая вначале движение электронов, будем считать тяжелые ядра неподвижными, что позволяет в гамильтониане (1.1) пренебречь слагаемым $\left(-\frac{\hbar^2}{2}\right)\sum_J \frac{1}{M_J}\nabla^2_{\boldsymbol{R}_J}$, связанным с кинетической энергией ядер. Волновая функция электронов φ, движущихся в поле неподвижных атомных ядер, удовлетворяет уравнению

$$\left\{-\frac{\hbar^2}{2m}\sum_i \nabla^2_{\boldsymbol{r}_i}+V(r, R)\right\}\varphi=\mathcal{E}\varphi. \qquad (1.2)$$

Теперь $\boldsymbol{R}_J$ — не переменные дифференциального уравнения, а параметры, определяющие потенциальное поле ядер. Очевидно,

что собственная функция и собственные значения уравнения (1.2) зависят от $\boldsymbol{R}_J$ как от параметров: $\varphi(\boldsymbol{r}, R)$ и $\mathscr{E}(R)$ [1]).

Попробуем представить полную волновую функцию системы электронов и ядер в виде

$$\Psi(\boldsymbol{r}, R) = \Phi(R)\,\varphi(\boldsymbol{r}, R). \tag{1.3}$$

Как мы увидим ниже, такое представление функции Ψ является приближенным, так как, строго говоря, отношение $\frac{\Psi(\boldsymbol{r}, R)}{\varphi(\boldsymbol{r}, R)}$ слабо зависит от $\boldsymbol{r}_i$. Функция $\Phi(R)$ окажется волновой функцией, описывающей движение ядер в среднем поле, создаваемом электронами. Подставив выражение (1.3) в уравнение (1.1), получим, учитывая (1.2),

$$-\frac{\hbar^2}{2}\sum_J \frac{1}{M_J}\left[\varphi\nabla^2_{R_J}\Phi + 2\left(\nabla_{R_J}\varphi\nabla_{R_J}\Phi\right) + \Phi\nabla^2_{R_J}\varphi\right] + \mathscr{E}(R)\,\varphi\Phi = W\varphi\Phi. \tag{1.4}$$

В самом деле, Φ от $\boldsymbol{r}_i$ не зависит, поэтому применение к произведению $\Phi\varphi$ оператора кинетической энергии электронов плюс V сразу дает, согласно уравнению (1.2), слагаемое $\mathscr{E}\varphi\Phi$.

С другой стороны, применение оператора $\nabla^2_{R_J}$ к произведению $\Phi\varphi$, в котором каждый множитель зависит от $\boldsymbol{R}_J$, дает (X — прямоугольная координата $\boldsymbol{R}_J$) выражение

$$\frac{\partial^2}{\partial X^2}(\Phi, \varphi) = \varphi\frac{\partial^2\Phi}{\partial X^2} + 2\frac{\partial\varphi}{\partial X}\frac{\partial\Phi}{\partial X} + \Phi\frac{\partial^2\varphi}{\partial X^2}.$$

Учитывая все три прямоугольные проекции $\boldsymbol{R}_J$, получим выражение, стоящее в квадратных скобках уравнения (1.4).

Не ограничивая общности, можно считать, что волновая функция электронов φ вещественна (мы исключаем наличие макроскопических токов в кристалле); тогда условие нормировки имеет вид

$$\int \varphi^2\,d\tau = 1, \tag{1.5}$$

где интегрирование ведется по всем координатам электронов по объему кристалла.

Из условия (1.5) следует:

$$\nabla_{R_J}\int \varphi^2\,d\tau = 2\int \varphi\nabla_{R_J}\varphi\,d\tau = 0. \tag{1.5а}$$

Умножим уравнение (1.4) на φ и проинтегрируем по τ. Используя (1.5) и (1.5а), получим

$$-\frac{\hbar^2}{2}\sum_J \frac{1}{M_J}\nabla^2_{R_J}\Phi + \left[\mathscr{E}(R) - \frac{\hbar^2}{2}\sum_J \frac{1}{M_J}\int \varphi\nabla^2_{R_J}\varphi\,d\tau\right]\Phi = W\Phi. \tag{1.6}$$

[1]) Учитывая вид $V(r, R)$, видно, что $\mathscr{E}(R)$ включает также кулоновскую энергию взаимодействия ядер.

Более тщательный анализ показывает, что в последнем выражении можно в квадратных скобках пренебречь $\sum_J$. Это связано с тем, что из-за малости отношений m/M_J волновая функция $\varphi(\boldsymbol{r}, R)$ слабо зависит от $\boldsymbol{R}_J$, так что можно пренебречь членами, содержащими $\nabla^2_{\boldsymbol{R}_J}\varphi$. Можно показать, что учет этих членов в уравнении (1.6) приводит при расчете различных физических величин к поправкам порядка $(m/M_J)^{1/4}$.

Таким образом, вместо (1.6) получим

$$\left[-\frac{\hbar^2}{2}\sum_J \frac{1}{M_J}\nabla^2_{\boldsymbol{R}_J} + \mathscr{E}(R)\right]\Phi = W\Phi. \qquad (1.7)$$

Мы видим, что функция Φ, определяемая дифференциальным уравнением (1.7), действительно зависит от переменных $\boldsymbol{R}_J$. Из (1.7) следует, что $\Phi(R)$ — волновая функция ядер, движущихся в поле с потенциальной энергией $\mathscr{E}(R)$, равной собственной энергии электронной системы при заданной конфигурации ядер и энергии кулоновского взаимодействия ядер.

Мы видим, что точная квантовомеханическая задача о поведении системы электронов и ядер распадается в адиабатическом приближении на две более простые: 1) задачу о движении электронов в поле неподвижных ядер (1.2) и 2) задачу о движении ядер в усредненном поле, создаваемом электронами $\mathscr{E}(R)$ [1]. Следует иметь в виду, что в некоторых явлениях, например при многофононных переходах электронов в примесных центрах, второе слагаемое в квадратной скобке уравнения (1.6), так называемый член неадиабатичности, играет существенную роль, являясь причиной этих переходов.

§ 2. Метод Хартри — Фока

1. В предыдущем параграфе мы видели, как на основе так называемого адиабатического приближения может быть упрощена общая квантовомеханическая задача о движении ядер и электронов в кристалле. Адиабатическое приближение, основанное на малости отношения m/M_J, позволяет свести вопрос о поведении электронно-ядерной системы твердого тела к задаче о движении электронов в поле неподвижных атомных ядер. Однако и в этом случае задача о движении совокупности всех электронов в кристалле остается чрезвычайно сложной и требует применения тех или иных приближенных методов. Одним из таких весьма эффективных методов, получивших преобладающее значение в электронной теории кристаллов, является метод

[1] Как отмечалось выше, $\mathscr{E}(R)$ включает также кулоновскую энергию взаимодействия ядер.

Хартри—Фока, позволяющий свести многоэлектронную задачу к одноэлектронной.

Рассмотрим систему N взаимодействующих друг с другом электронов в поле неподвижных ядер. Уравнение Шредингера для стационарных состояний этой системы имеет вид

$$\hat{\mathcal{H}}\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \mathcal{E}\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) \quad (2.1)$$

с гамильтонианом

$$\hat{\mathcal{H}} = \sum_{i}^{1, N} \hat{\mathcal{H}}_i + \frac{1}{2}\sum_{ij}^{1, N}{}' \frac{e^2}{r_{ij}}, \quad (2.1а)$$

где

$$\hat{\mathcal{H}}_i = -\frac{\hbar^2}{2m}\nabla_i^2 + V(\boldsymbol{r}_i). \quad (2.1б)$$

Здесь $\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}$; $V(\boldsymbol{r}_i)$ — потенциальная энергия i-го электрона в поле ядер, штрих у суммы в выражении (2.1а) указывает, что в потенциальной энергии кулоновского взаимодействия электронов надо опустить члены с $i = j$.

Метод Хартри—Фока основан на идее замены в гамильтониане $\hat{\mathcal{H}}$ потенциальной энергии взаимодействия электронов некоторым эффективным внешним полем $\mathcal{U}_{\text{eff}}(\boldsymbol{r})$, в котором каждый электрон движется независимо. Поле $\mathcal{U}_{\text{eff}}$ должно наилучшим способом описывать усредненное действие всех остальных электронов на данный. Гамильтониан системы равен теперь сумме гамильтонианов, каждый из которых зависит только от координат одного электрона, т. е.

$$\hat{\mathcal{H}} = \sum_{i=1}^{N} \hat{\mathcal{H}}'_i, \quad (2.2)$$

где

$$\hat{\mathcal{H}}'_i = -\frac{\hbar^2}{2m}\nabla_i^2 + V(\boldsymbol{r}_i) + \mathcal{U}_{\text{eff}}(r_i). \quad (2.2а)$$

Ниже будет показано, как наилучшим образом выбирать поле $\mathcal{U}_{\text{eff}}$. Нетрудно показать, что уравнение (2.1) с гамильтонианом (2.2) имеет решение:

$$\psi_*(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \psi_{n_1}(\boldsymbol{r}_1)\psi_{n_2}(\boldsymbol{r}_2)\ldots\psi_{n_N}(\boldsymbol{r}_N). \quad (2.3)$$

Индекс n_i у функции ψ_{n_i} обозначает три квантовых числа, характеризующих квантовое состояние i-го электрона, определяемое уравнением

$$\hat{\mathcal{H}}'_i\psi_{n_i}(\boldsymbol{r}_i) = \mathcal{E}_{n_i}\psi_{n_i}(\boldsymbol{r}_i), \quad (2.3а)$$

где $\mathcal{E}_{n_i}$ — соответствующее ему собственное значение энергии.

При этом полная энергия системы в приближении (2.2) равна

$$\mathcal{E} = \sum_{i=1}^{N} \mathcal{E}_{n_i}. \tag{2.3б}$$

В самом деле, подставим выражение (2.3) в уравнение (2.1) при условии, что $\hat{\mathcal{H}}$ равно сумме (2.2). Имеем

$$\sum_i \hat{\mathcal{H}}_i' [\psi_{n_1}(\boldsymbol{r}_1) \ldots \psi_{n_i}(\boldsymbol{r}_i) \ldots \psi_{n_N}(\boldsymbol{r}_N)] = \mathcal{E} [\psi_{n_1}(\boldsymbol{r}_1) \ldots \psi_{n_N}(\boldsymbol{r}_N)]$$

или

$$\sum_i [\psi_{n_1}(\boldsymbol{r}_1) \ldots \psi_{n_{i-1}}(\boldsymbol{r}_{i-1})\, \psi_{n_{i+1}}(\boldsymbol{r}_{i+1}) \ldots \psi_{n_N}(\boldsymbol{r}_N)]\, \hat{\mathcal{H}}_i' \psi_{n_i}(\boldsymbol{r}_i) = \mathcal{E} [\psi_{n_1}(\boldsymbol{r}_1) \ldots \psi_{n_N}(\boldsymbol{r}_N)],$$

так как оператор $\hat{\mathcal{H}}_i'$ действует только на координаты i-го электрона. Используя (2.3а), получим

$$\sum_i [\psi_{n_1}(\boldsymbol{r}_1) \ldots \psi_{n_N}(\boldsymbol{r}_N)]\, \mathcal{E}_{n_i} = \mathcal{E} [\psi_{n_1}(\boldsymbol{r}_1) \ldots \psi_{n_N}(\boldsymbol{r}_N)].$$

Сокращая обе части равенства на $[\psi_{n_1}(\boldsymbol{r}_1) \ldots \psi_{n_N}(\boldsymbol{r}_N)]$, получим выражение (2.3б).

Так как $|\psi_{n_i}(\boldsymbol{r}_i)^2|$ — плотность вероятности нахождения i-го электрона в точке пространства $\boldsymbol{r}_i$, то мультипликативный характер решения (2.3) характеризует с точки зрения теоремы об умножении вероятностей независимый характер движения во внешнем поле $V + \mathcal{U}_{\text{eff}}$ не взаимодействующих друг с другом электронов.

2. Поставим теперь задачу наилучшего определения $\mathcal{U}_{\text{eff}}(\boldsymbol{r})$ в приближенном гамильтониане (2.2а). Как мы увидим, это можно будет сделать на основе некоторой самосогласованной процедуры.

Для того чтобы дальнейшие выкладки были менее громоздки, проведем их для системы из $N = 2$ электронов, а затем обобщим результаты на произвольное число электронов N.

Для двух электронов выражение (2.3) имеет вид

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_{n_1}(\boldsymbol{r}_1)\, \psi_{n_2}(\boldsymbol{r}_2).$$

Мы должны теперь учесть, что состояние электрона характеризуется наряду с тремя пространственными координатами $x, y, z\ (\equiv \boldsymbol{r})$ значением проекции собственного (внутреннего) момента количества движения (спина) S_z на некоторое заданное направление (например, ось z). Теория и опыт показывают, что S_z принимает для электрона только два значения $+\hbar/2$ и $-\hbar/2$; если мы положим $S_z = s\hbar$, то спиновая координата $s = +1/2$ или $s = -1/2$. В соответствии с этим мы введем спиновые функции $\nu_i(s)$ $(i = 1, 2)$, где индекс i характеризует спиновое состояние:

$i=1$ для $S_z=\hbar/2$ и $i=2$ для $S_z=-\hbar/2$, так что

$$\nu_1(^1/_2)=1, \quad \nu_1(-^1/_2)=0, \quad \nu_2(^1/_2)=0, \quad \nu_2(-^1/_2)=1. \quad (2.4)$$

При таком определении спиновые функции ортонормированы, т. е.

$$\sum_{s=\pm^1/_2} \nu_i^*(s)\,\nu_k(s)=\delta_{ik}. \quad (2.5)$$

Если не учитывать взаимодействие магнитного момента электрона (связанного со спином) с магнитным полем, создаваемым его орбитальным движением, то полная одноэлектронная волновая функция l-го электрона, находящегося в $(n_i, k)\equiv j$-квантовом состоянии, равна

$$\psi_{n_i}(\boldsymbol{r}_l)\,\nu_k(s_l)=\varphi_j(l). \quad (2.6)$$

Аргумент l у функции φ_j обозначает совокупность четырех координат l-го электрона: трех пространственных x_l, y_l, z_l и спиновой s_l. Предполагая волновые функции ψ_{n_i} ортонормированными, получим

$$\int \varphi_j^*(l)\,\varphi_{j'}(l)\,d\tau_l=\int d\boldsymbol{r}_l \sum_s \psi_{n_i}^*(\boldsymbol{r}_l)\,\psi_{n_i'}(\boldsymbol{r}_l)\,\nu_k^*(s)\,\nu_{k'}(s)=$$

$$=\delta_{n_i n_i'}\delta_{kk'}=\delta_{jj'}. \quad (2.7)$$

Здесь символ $\int d\tau_l$ означает интегрирование по пространственным координатам x_l, y_l, z_l и суммирование по спиновой координате s.

Согласно принципу Паули полная волновая функция системы должна быть антисимметрична, т. е. при перемене пары электронов местами (при перестановке их четырех координат) должна менять свой знак. Для двух электронов ($N=2$) такая волновая функция имеет вид

$$\Phi(1,2)=\frac{1}{\sqrt{2}}\{\varphi_1(1)\,\varphi_2(2)-\varphi_1(2)\,\varphi_2(1)\}=\frac{1}{\sqrt{2!}}\begin{vmatrix}\varphi_1(1) & \varphi_1(2)\\ \varphi_2(1) & \varphi_2(2)\end{vmatrix}. \quad (2.8)$$

В самом деле, $\Phi(2.1)=-\Phi(1.2)$. Множитель $1/\sqrt{2}$ введен для нормировки волновой функции системы. Используя соотношение (2.7), легко показать, что

$$\int \Phi^*(1,2)\,\Phi(1,2)\,d\tau_1\,d\tau_2=1. \quad (2.8а)$$

Антисимметричная форма выражения (2.8) автоматически обеспечивает требование принципа Паули, согласно которому в одном квантовом состоянии не может находиться больше одного электрона. В самом деле, если $\varphi_1=\varphi_2$, то $\Phi=0$.

Обобщая определитель (2.8) на N электронов, получим правильную антисимметричную волновую функцию для системы из N

электронов:

$$\Phi(1, 2, 3, \ldots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(1) & \varphi_1(2) & \ldots & \varphi_1(N) \\ \varphi_2(1) & \varphi_2(2) & \ldots & \varphi_2(N) \\ \ldots & \ldots & \ldots & \ldots \\ \varphi_N(1) & \varphi_N(2) & \ldots & \varphi_N(N) \end{vmatrix}, \quad (2.8б)$$

удовлетворяющую принципу Паули.

Если обменять в этом выражении координаты пары электронов, например $1 \rightleftarrows 2$, то это эквивалентно перестановке двух столбцов определителя, а при этом он изменит свой знак, т. е. $\Phi(1, 2, 3, \ldots, N) = -\Phi(2, 1, 3, \ldots, N)$. Если считать два квантовых состояния совпадающими, например $\varphi_1 = \varphi_2$, то две строки определителя равны, а в этом случае $\Phi = 0$. Множитель $1/\sqrt{N!}$ обеспечивает нормировку функции $\Phi(1, 2, \ldots, N)$ при условии (2.7).

Используя точный гамильтониан системы (2.1а) и волновую функцию (2.8б), вычислим полную энергию системы (Приложение 8, п. 1):

$$\mathcal{E} = \int \Phi^* \hat{\mathcal{H}} \Phi \, d\tau_1 \, d\tau_2 \ldots d\tau_N =$$

$$= \sum_i^{1, N} \int \varphi_i^*(1) \hat{\mathcal{H}}_1 \varphi_i(1) \, d\tau_1 + \frac{1}{2} \sum_{i, j}^{1, N}{}' \int |\varphi_i(1)|^2 \frac{e^2}{r_{12}} |\varphi_j(2)|^2 \, d\tau_1 \, d\tau_2 -$$

$$- \frac{1}{2} \sum_{i, j}^{1, N}{}' \int \varphi_i^*(1) \varphi_j(1) \frac{e^2}{r_{12}} \varphi_i(2) \varphi_j^*(2) \, d\tau_1 \, d\tau_2. \quad (2.9)$$

Заметим, что штрихи у второй и третьей сумм могут быть опущены, так как члены с $i = j$ сокращаются.

Так как оператор $\hat{\mathcal{H}}_1$ и e^2/r_{12} не зависят от спиновой переменной s_1, то суммирование по ней может быть выполнено независимо от интегрирования по пространственным координатам. Очевидно, что в тех случаях, когда суммирование по s_1 ведется для пары функций φ_i с одинаковыми индексами, результат, согласно условию (2.5), равен единице. Для интегралов последней суммы в выражении (2.9) результат равен единице только при суммировании по электронам с параллельными спинами, в противном случае суммирование дает нуль.

Таким образом, выражение (2.9) может быть записано в виде

$$\mathcal{E} = \sum_i^{1, N} \int \psi_{n_i}^*(\boldsymbol{r}_1) \hat{\mathcal{H}}_1 \psi_{n_i}(\boldsymbol{r}_1) \, d\boldsymbol{r}_1 +$$

$$+ \frac{1}{2} \sum_{i, j}^{1, N} \int |\psi_{n_i}(\boldsymbol{r}_1)|^2 \frac{e^2}{r_{12}} |\psi_{n_j}(\boldsymbol{r}_2)|^2 \, d\boldsymbol{r}_1 \, d\boldsymbol{r}_2 -$$

$$- \frac{1}{2} \sum_{i, j}^{1, N} \int \psi_{n_i}^*(\boldsymbol{r}_1) \psi_{n_j}(\boldsymbol{r}_1) \frac{e^2}{r_{12}} \psi_{n_i}(\boldsymbol{r}_2) \psi_{n_j}^*(\boldsymbol{r}_2) \, d\boldsymbol{r}_1 \, d\boldsymbol{r}_2, \quad (2.9а)$$

где в последней сумме суммирование ведется по состояниям электронов с параллельными спинами.

Так как электронные функции ψ_{n_i} не являются собственными функциями гамильтониана $\hat{\mathcal{H}}_1$, то первая сумма не есть сумма энергии невзаимодействующих электронов. Интегралы, входящие под знак второй суммы (положительной), называются *кулоновскими* — они равны потенциальной энергии взаимодействия двух зарядов, распределенных с плотностями: $-e|\psi_{n_i}(\boldsymbol{r}_1)|^2$ и $-e|\psi_{n_j}(\boldsymbol{r}_2)|^2$. Интегралы, входящие под знак последней суммы, называются *обменными* — они не имеют классического электростатического аналога, им соответствует кулоновское взаимодействие с «комплексными плотностями заряда»:

$$e\psi^*_{n_i}(\boldsymbol{r}_1)\psi_{n_j}(\boldsymbol{r}_1) \quad \text{и} \quad -e\psi^*_{n_j}(\boldsymbol{r}_2)\psi_{n_i}(\boldsymbol{r}_2).$$

Оба электрона с параллельными спинами находятся частью в n_i-м, частью в n_j-м состояниях, они как бы «обмениваются» местами.

Остается открытым вопрос, как наилучшим образом выбрать одноэлектронные волновые функции ψ_{n_i}. Для определения наилучших одноэлектронных волновых функций $\psi_{n_i}(\boldsymbol{r})$ потребуем, чтобы энергия системы $\mathcal{E}$ была минимальна по отношению к произвольным малым изменениям (вариациям) функций $\psi_{n_i} \to \psi_{n_i} + \delta\psi_{n_i}$. Так как функции ψ_{n_i} комплексны (т. е. эквивалентны двум вещественным функциям), то необходимо независимо варьировать ψ_{n_i} и $\psi^*_{n_i}$. Однако уравнения, которые получаются при вариации $\psi^*_{n_i}$, комплексно сопряжены с теми, которые получаются при вариации ψ_{n_i}, поэтому мы ограничимся вариацией одной из функций, например: $\psi^*_{n_i} \to \psi^*_{n_i} + \delta\psi^*_{n_i}$. Мы потребуем выполнения условия ортонормированности и по отношению к варьированным функциям:

$$\int (\psi^*_{n_i} + \delta\psi^*_{n_i})\psi^*_{n_j}\,d\boldsymbol{r} = \delta_{n_i n_j}. \tag{2.10}$$

Учитывая условие (2.7), получим

$$\int \delta\psi^*_{n_i}\psi_{n_j}\,d\boldsymbol{r} = 0 \tag{2.10а}$$

для всех i и j.

Варьируя в правой части уравнения (2.9а) функции $\psi^*_{n_i}$, вычислим соответствующее изменение энергии системы $\delta\mathcal{E}$. Приравнивая $\delta\mathcal{E}$ нулю при дополнительном условии (2.10а), которое учитывается методом неопределенных множителей Лагранжа, получим следующее уравнение для определения функции ψ_{n_i}

(Приложение 8, п. 2):

$$\hat{\mathcal{H}}_1\psi_{n_i}(\boldsymbol{r}_1)+\left[\sum_j{}'\int\frac{e^2\,|\,\psi_{n_j}(\boldsymbol{r}_2)\,|^2}{r_{12}}\,d\boldsymbol{r}_2\right]\psi_{n_i}(\boldsymbol{r}_1)-$$

$$-\sum_j{}'\int\frac{e^2\psi_{n_i}(\boldsymbol{r}_2)\,\psi^*_{n_j}(\boldsymbol{r}_2)}{r_{12}}\,d\boldsymbol{r}_2\psi_{n_j}(\boldsymbol{r}_1)-\mathcal{E}_{n_i}\psi_{n_i}(\boldsymbol{r}_1)=0\ (i=1,2,\ldots).\quad(2.11)$$

Здесь $\mathcal{E}_{n_i}$ — неопределенный множитель Лагранжа, играющий роль собственного значения энергии одноэлектронного состояния. Вторую сумму в выражении (2.11) можно формально представить в виде оператора, действующего на ψ_{n_i}, и записать уравнение (2.11) в виде

$$\left[\hat{\mathcal{H}}_1+\sum_j{}'\int\frac{e^2\,|\,\psi_{n_j}(\boldsymbol{r}_2)\,|^2}{r_{12}}\,d\boldsymbol{r}_2-\right.$$

$$\left.-\sum_j{}'\frac{\psi_{n_j}(\boldsymbol{r}_1)}{\psi_{n_i}(\boldsymbol{r}_1)}\int e^2\psi_{n_i}(\boldsymbol{r}_2)\,\psi^*_{n_i}(\boldsymbol{r}_2)\,d\boldsymbol{r}_2\right]\psi_{n_i}(\boldsymbol{r}_1)=\mathcal{E}_{n_i}\psi_{n_i}(\boldsymbol{r}_1).\quad(2.11а)$$

Из сравнения этого выражения с (2.3а) видно, что

$$\mathcal{U}_{\text{eff}}(\boldsymbol{r}_1)=\sum_j{}'\int\frac{e^2\,|\,\psi_{n_j}(\boldsymbol{r}_2)\,|^2}{r_{12}}\,d\boldsymbol{r}_2-$$

$$-\sum_j{}'\frac{\psi_{n_j}(\boldsymbol{r}_1)}{\psi_{n_i}(\boldsymbol{r}_1)}\int\frac{e^2\psi_{n_i}(\boldsymbol{r}_2)\,\psi_{n_j}(\boldsymbol{r}_2)}{r_{12}}\,d\boldsymbol{r}_2,\quad(2.11б)$$

где во второй сумме суммирование ведется по электронам со спинами, параллельными спину электрона в состоянии n_i.

Уравнения (2.11а) называются *уравнениями самосогласованного поля Хартри—Фока*. Хартри задавал волновую функцию системы в виде (2.3), не учитывающей принцип Паули; он формулировал уравнения (2.11а) в виде, в котором отсутствовала вторая сумма в квадратных скобках, что соответствует пренебрежению энергией обменного взаимодействия. Это пренебрежение приводит к ошибкам в 10—20% при вычислении плотности электронного облака в многоэлектронных атомах. Строгая теория самосогласованного поля многоэлектронных систем была развита В. А. Фоком (1930).

Так как $\mathcal{U}_{\text{eff}}(\boldsymbol{r}_1)$ само зависит от неизвестных функций ψ_{n_i}, то уравнение (2.11а) представляет собой систему нелинейных интегродифференциальных уравнений для функций ψ_{n_i}. Можно представить себе следующую процедуру их решения. Зададим в нулевом приближении некоторые одноэлектронные волновые функции ψ_{n_i} и вычислим с ними эффективное поле $\mathcal{U}_{\text{eff}}$. После этого можно из системы уравнений (2.11а) определить функции ψ_{n_i} в следующем, первом приближении. Определив посредством

новых ψ_{n_i} поле $\mathcal{U}_{\text{eff}}$ в первом приближении, можно, решая уравнения (2.11а), определить ψ_{n_i} во втором приближении и т. д. Если функции, подставляемые в $\mathcal{U}_{\text{eff}}$ будут совпадать с теми, которые будут найдены из решения уравнения (2.11а), то можно считать, что определены самосогласованные решения уравнений Хартри—Фока. Конечно, фактическое проведение такой операции представляет значительные трудности.

§ 3. Электрон в периодическом поле

1. Как показывают теория и опыт, многоэлектронная задача в кристалле во многих случаях может быть с достаточным приближением, рассмотрена как одноэлектронная. Другими словами, электроны в кристалле могут быть с хорошим приближением описаны уравнениями Хартри—Фока. Как должно быть в этом случае выбрано эффективное поле $\mathcal{U}_{\text{eff}}(\boldsymbol{r})$ в (2.2а)?

Симметрия кристалла подсказывает, что $\mathcal{U}_{\text{eff}}(\boldsymbol{r})$ должно обладать периодичностью кристалла.

В гл. II, § 9, п. 2 было показано, что волновая функция электрона в периодическом поле имеет вид (II.9.13)

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = u_{\boldsymbol{k}}(\boldsymbol{r})\, e^{i\boldsymbol{k}\boldsymbol{r}}, \tag{3.1}$$

где $\boldsymbol{k}$—волновой вектор электрона, а амплитудная функция

$$u_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{a}_n) = u_{\boldsymbol{k}}(\boldsymbol{r}), \tag{3.1а}$$

т. е. обладает периодичностью кристаллической решетки. Можно показать (Приложение 9), что блоховские волновые функции (3.1) подставленные в выражение (2.11б), действительно приводят к эффективному полю $\mathcal{U}_{\text{eff}}(\boldsymbol{r})$, обладающему периодичностью кристалла, т. е. что решение (3.1) самосогласовано.

Волновой вектор электрона может быть представлен в виде (II.9.8):

$$\boldsymbol{k} = \frac{g_1}{G}\boldsymbol{b}_1 + \frac{g_2}{G}\boldsymbol{b}_2 + \frac{g_3}{G}\boldsymbol{b}_3 \qquad (g_i = 0, 1, 2, \ldots, G-1), \tag{3.2}$$

где G—большое (нечетное) число, а $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ и $\boldsymbol{b}_3$—основные векторы обратной решетки. Таким образом, $\boldsymbol{k}$ принимает G^3 квазидискретных значений.

К этому же выражению для $\boldsymbol{k}$ мы придем, если выделим основную область кристалла в форме параллелепипеда с ребрами $G\boldsymbol{a}_1$, $G\boldsymbol{a}_2$, $G\boldsymbol{a}_3$ и объемом $V = G^3\Omega_0$ ($\Omega_0 = (\boldsymbol{a}_1[\boldsymbol{a}_2, \boldsymbol{a}_3])$—объем элементарной ячейки кристалла), и потребуем, чтобы волновая функция (3.1) не менялась при смещении на вектор $G\boldsymbol{a}_i$ ($i = 1, 2, 3$) (условия *цикличности Борна—Кармана*; см. гл. III, § 5, п. 6). Физически различные (неэквивалентные) значения $\boldsymbol{k}$ лежат

в пределах (II.9.10):

$$-\pi < \boldsymbol{k}\boldsymbol{a}_i < \pi \quad (i = 1, 2, 3). \tag{3.3}$$

Подставляя сюда (3.2), получим (II.9.10а)

$$-\frac{G}{2} < g_i < \frac{G}{2}, \tag{3.3а}$$

так что в области изменения вектора $\boldsymbol{k}$ (3.3) он принимает G^3 квазидискретных значений.

В качестве наиболее удобной области неэквивалентных (разных) значений волнового вектора $\boldsymbol{k}$, мы выберем бриллюэновскую зону, так как она определена в гл. II, § 9, п. 1.

Условия (3.2), (3.3) и (3.3а) совпадают с условиями (III.5.9), (III.5.24), (III.5.25) для волнового вектора $\boldsymbol{q}$, что вполне естественно, так как само существование волнового вектора и его свойства обусловлены наличием трансляционной симметрии кристаллической решетки.

Если блоховская волновая функция (3.1) нормирована на основную область кристалла, то

$$\int\limits_V \psi_{\boldsymbol{k}}^*(\boldsymbol{r})\,\psi_{\boldsymbol{k}}(\boldsymbol{r})\,d^3r = \int\limits_{G^3\Omega_0} |u_{\boldsymbol{k}}(\boldsymbol{r})|^2\,d^3r = 1. \tag{3.4}$$

Так как $|u_{\boldsymbol{k}}(\boldsymbol{r})|^2$ периодично с периодами основных векторов решетки, то из (3.4) следует

$$G^3 \int\limits_{\Omega_0} |u_{\boldsymbol{k}}(\boldsymbol{r})|^2\,d^3r = 1, \tag{3.4а}$$

где интегрирование ведется по элементарной ячейке кристалла. В некоторых случаях представляется целесообразным ввести в блоховскую функцию в явном виде множитель $G^{-3/2}$, т. е. положить

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{G^{3/2}}\,u_{\boldsymbol{k}}(\boldsymbol{r})\,e^{i\boldsymbol{k}\boldsymbol{r}}. \tag{3.5}$$

Тогда вместо (3.4а) получим

$$\int\limits_{\Omega_0} |u_{\boldsymbol{k}}(\boldsymbol{r})|^2\,d^3r = 1, \tag{3.6}$$

и в этом случае среднее значение

$$\overline{|u_{\boldsymbol{k}}|} = 1/\sqrt{\Omega_0}. \tag{3.6а}$$

Подставляя блоховскую функцию (3.1) в уравнение Шредингера

$$-\frac{\hbar^2}{2m}\nabla^2\psi_{\boldsymbol{k}} + V(\boldsymbol{r})\,\psi_{\boldsymbol{k}} = \varepsilon_{\boldsymbol{k}}\psi_{\boldsymbol{k}}, \tag{3.7}$$

где $V(\boldsymbol{r})$ — периодический потенциал, действующий на электрон в кристалле, получим, после сокращения на множитель $\exp(i\boldsymbol{k}\boldsymbol{r})$, следующее уравнение для $u_{\boldsymbol{k}}(\boldsymbol{r})$ (Приложение 10):

$$-\frac{\hbar^2}{2m}\nabla^2 u_{\boldsymbol{k}} + V(\boldsymbol{r})\,u_{\boldsymbol{k}} - \frac{i\hbar^2}{m}(\boldsymbol{k}\cdot\nabla u_{\boldsymbol{k}}) = \left(\varepsilon_{\boldsymbol{k}} - \frac{\hbar^2 k^2}{2m}\right)u_{\boldsymbol{k}}. \qquad (3.8)$$

Здесь $\nabla u_{\boldsymbol{k}} \equiv \operatorname{grad} u_{\boldsymbol{k}}$.

Для $\boldsymbol{k}=0$ уравнение (3.8) переходит в уравнение

$$-\frac{\hbar^2}{2m}\nabla^2 u_0 + V(\boldsymbol{r})\,u_0 = \varepsilon_0 u_0, \qquad (3.8а)$$

совпадающее с уравнением (3.7) для $\psi_{\boldsymbol{k}}$.

2. Изучение разных видов периодических полей (одномерные модели, случаи слабой и сильной связи) и общие соображения показывают, что не все значения энергии электрона $\varepsilon_{\boldsymbol{k}}$ в (3.7) являются разрешенными. Энергетический спектр электрона в периодическом поле, в общем случае, распадается на *разрешенные* и *запрещенные полосы* или *зоны энергии* (не путать с бриллюэновскими зонами).

Другими словами, зависимость энергии электрона в периодическом поле от волнового вектора $\boldsymbol{k}$ (в пределах первой бриллюэновской зоны) не является однозначной функцией, а состоит из множества $\varepsilon_n(\boldsymbol{k})$, где n — номер зоны (полосы) разрешенной энергии. Ситуация аналогична образованию ветвей колебаний $\omega_j(\boldsymbol{q})$ $(j=1, 2, \ldots, 3s)$ (см. гл. III, § 5, п. 2) с тем только отличием, что число разрешенных зон энергии $\varepsilon_n(\boldsymbol{k})$ бесконечно велико.

Волновая функция электрона (3.1) будет зависеть не только от волнового вектора $\boldsymbol{k}$, но и от номера зоны n, т. е.

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = u_{n\boldsymbol{k}}(\boldsymbol{r})\exp(i\boldsymbol{k}\boldsymbol{r}).$$

Из физической эквивалентности волновых векторов $\boldsymbol{k}$ и $\boldsymbol{k}' = \boldsymbol{k}+\boldsymbol{b}_g$ ($\boldsymbol{b}_g$ — вектор обратной решетки) следует, что все величины, зависящие от $\boldsymbol{k}$, должны быть периодичны с периодами основных векторов обратной решетки $\boldsymbol{b}_i$ $(i=1, 2, 3)$.

В частности, для энергии $\varepsilon_n(\boldsymbol{k})$ имеем

$$\varepsilon_n(\boldsymbol{k}+\boldsymbol{b}_g) = \varepsilon_n(\boldsymbol{k}), \qquad (3.9)$$

что означает, что $\varepsilon_n(\boldsymbol{k})$ может быть разложено в трехмерный ряд Фурье в $\boldsymbol{k}$-пространстве:

$$\varepsilon_n(\boldsymbol{k}) = \sum_{\boldsymbol{l}} c_{\boldsymbol{l}} e^{i\boldsymbol{k}\boldsymbol{a}_{\boldsymbol{l}}}, \qquad (3.10)$$

где $\boldsymbol{l} \equiv \{l_1, l_2, l_3\}$ — целочисленные индексы вектора прямой решетки (I.3.1); в самом деле, заменяя в (3.10) $\boldsymbol{k}$ на $\boldsymbol{k}+\boldsymbol{b}_g$ подтвердим (3.9) ($\exp(i\boldsymbol{b}_g\boldsymbol{a}_l) = \exp(2\pi i \times$ целое число$) = 1$). Разложение (3.10) аналогично разложению (I.3.6) в $\boldsymbol{r}$-пространстве.

Вычислим действие оператора $\exp(\boldsymbol{a}_l\cdot\nabla)$ на блоховскую волновую функцию $\psi_{n\boldsymbol{k}}(\boldsymbol{r})$. Разлагая экспоненту в ряд, получим

$$\exp(\boldsymbol{a}_l\cdot\nabla)\,\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = [1+\boldsymbol{a}_l\cdot\nabla+{}^1/_2(\boldsymbol{a}_l\cdot\nabla)^2+\ldots]\,\psi_{n\boldsymbol{k}}(\boldsymbol{r}) =$$

$$= \psi_{n\boldsymbol{k}}(\boldsymbol{r})+\boldsymbol{a}_l\nabla\psi_{n\boldsymbol{k}}(\boldsymbol{r})+\frac{1}{2}\sum_{\alpha,\,\beta}^{1,\,2,\,3}\boldsymbol{a}_{l\alpha}\boldsymbol{a}_{l\beta}\frac{\partial}{\partial x_\alpha}\frac{\partial}{\partial x_\beta}[\psi_{n\boldsymbol{k}}(\boldsymbol{r})]+\ldots =$$

$$=\psi_{n\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{a}_l), \quad (3.11)$$

где полученный ряд можно рассматривать как ряд Тейлора по степеням прямоугольных составляющих вектора $\boldsymbol{a}_l$.

Подставляя в (3.10) $-i\nabla$ вместо $\boldsymbol{k}$ и используя (3.11), получим

$$\varepsilon_n(-i\nabla)\,\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_l c_l e^{\boldsymbol{a}_l\cdot\nabla}\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_l c_l\psi_{n\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{a}_l) =$$

$$= \sum_l c_l e^{i\boldsymbol{k}\boldsymbol{a}_l}\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \varepsilon_n(\boldsymbol{k})\,\psi_{n\boldsymbol{k}}(\boldsymbol{r}). \quad (3.12)$$

Соотношение (3.12) получило название *теоремы Ванье*; при ее выводе предполагалось, что энергия $\varepsilon_n(\boldsymbol{k})$ при данном $\boldsymbol{k}$ не вырождена.

Пусть на электрон проводимости в кристалле, помимо периодического поля $V(\boldsymbol{r})$, действует дополнительное поле $\mathcal{U}(\boldsymbol{r})$ (например, внешнее электрическое поле или потенциал примесного иона); тогда уравнение Шредингера для электрона имеет вид

$$\hat{\mathcal{H}}\psi(\boldsymbol{r}) = \left[-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r})+\mathcal{U}(\boldsymbol{r})\right]\psi(\boldsymbol{r}) = \varepsilon\psi(\boldsymbol{r}). \quad (3.13)$$

Разложим волновую функцию $\psi(\boldsymbol{r})$ по полной системе блоховских функций $\psi_{n\boldsymbol{k}}(\boldsymbol{r})$, удовлетворяющих уравнению (3.7):

$$\psi(\boldsymbol{r}) = \sum_n\sum_{\boldsymbol{k}} c_{n\boldsymbol{k}}\psi_{n\boldsymbol{k}}(\boldsymbol{r}). \quad (3.14)$$

Здесь $c_{n\boldsymbol{k}}$ — коэффициенты разложения.

Подставляя (3.14) в (3.13) и учитывая, что $\psi_{n\boldsymbol{k}}(\boldsymbol{r})$ — собственные функции уравнения (3.7), получим

$$\hat{\mathcal{H}}\psi(\boldsymbol{r}) = \sum_n\sum_{\boldsymbol{k}} c_{n\boldsymbol{k}}[\varepsilon_n(\boldsymbol{k})+\mathcal{U}(\boldsymbol{r})]\,\psi_{n\boldsymbol{k}} = \varepsilon\psi(\boldsymbol{r}). \quad (3.15)$$

Используя (3.12), получим

$$\hat{\mathcal{H}}\psi(\boldsymbol{r}) = \sum_n[\varepsilon_n(-i\nabla)+\mathcal{U}(\boldsymbol{r})]\sum_{\boldsymbol{k}} c_{n\boldsymbol{k}}\psi_{n\boldsymbol{k}} = \varepsilon\psi(\boldsymbol{r}). \quad (3.16)$$

Если с достаточным приближением можно использовать блоховские функции одной зоны, например n-й, то

$$[\varepsilon_n(-i\nabla)+\mathcal{U}(\boldsymbol{r})]\,\psi(\boldsymbol{r}) = \varepsilon\psi(\boldsymbol{r}). \quad (3.17)$$

В этом приближении уравнение Шредингера (3.13) заменяется уравнением (3.17), в котором уже не фигурирует периодический

потенциал $V(\boldsymbol{r})$. На первый взгляд практическая польза от этого невелика, так как в уравнении (3.17) надо знать $\varepsilon_n(\boldsymbol{k})$ во всей зоне Бриллюэна, а для этого надо определить собственные значения энергии в уравнении (3.13). Можно, однако, пытаться использовать приближенные выражения для $\varepsilon_n(\boldsymbol{k})$.

3. Если в точке $\boldsymbol{k}=\boldsymbol{k}_0$ энергия $\varepsilon_n(\boldsymbol{k})$ обладает минимумом или максимумом, то ее можно вблизи этой точки разложить в ряд

$$\varepsilon_n(\boldsymbol{k})=\varepsilon_n(\boldsymbol{k}_0)+\frac{1}{2}\sum_{\alpha,\beta}\left(\frac{\partial^2\varepsilon_n(\boldsymbol{k})}{\partial k_\alpha \partial k_\beta}\right)_{k=k_0}(k_\alpha-k_{\alpha 0})(k_\beta-k_{\beta 0}), \quad (3.18)$$

где k_α и k_β — прямоугольные составляющие вектора $\boldsymbol{k}$. Мы ограничились в (3.18) разложением до квадратичных членов и учли, что в точке экстремума первые производные $(\partial\varepsilon_n/\partial k_\alpha)_{\boldsymbol{k}=\boldsymbol{k}_0}=0$. Так как энергия ε_n — скаляр, а $\boldsymbol{k}$ — вектор, то величины $(\partial^2\varepsilon_n/\partial k_\alpha \partial k_\beta)_{k_0}$ — компоненты тензора второго ранга (Приложение 11). Приводя этот тензор к главным осям и перенося начало отсчета энергии и волнового вектора в точку экстремума ($\varepsilon_n(\boldsymbol{k}_0)=0$, $\boldsymbol{k}_0=0$), получим

$$\varepsilon_n(\boldsymbol{k})=\frac{1}{2}\sum_{\alpha}\left(\frac{\partial^2\varepsilon_n}{\partial k_\alpha^2}\right)_0 k_\alpha^2. \quad (3.18\text{а})$$

Для того чтобы максимально приблизить описание движения электрона в периодическом поле к движению свободного электрона, введем тензор *обратной эффективной массы*

$$\frac{1}{m_{\alpha\beta}}=\frac{1}{\hbar^2}\left(\frac{\partial^2\varepsilon_n}{\partial k_\alpha \partial k_\beta}\right)_0, \quad (3.19)$$

компоненты которого в главных осях равны

$$\frac{1}{m_\alpha}=\frac{1}{\hbar^2}\left(\frac{\partial^2\varepsilon_n}{\partial k_\alpha^2}\right)_0. \quad (3.19\text{а})$$

Введем *квазиимпульс*

$$\boldsymbol{p}=\hbar\boldsymbol{k}, \quad (3.20)$$

отличающийся от волнового вектора $\boldsymbol{k}$ только постоянным множителем $\hbar$ и поэтому обладающий свойствами (3.2)—(3.3а).

Заметим, что тензор обратной эффективной массы имеет размерность (масса)$^{-1}$, а квазиимпульс — размерность импульса. Из выражений (3.18а), (3.19а) и (3.20) следует, что

$$\varepsilon_n(\boldsymbol{k})=\sum_\alpha\frac{\hbar^2k_\alpha^2}{2m_\alpha}=\sum_\alpha\frac{p_\alpha^2}{2m_\alpha}, \quad (3.21)$$

т. е. имеет вид кинетической энергии электрона с различными массами по осям x, y и z. Величины m_α, имеющие размерность массы, не являются компонентами тензора (так как величины, обратные компонентам тензора, вообще говоря, не образуют

тензор), однако для сокращения их часто называют *тензором эффективной массы*. Следует иметь в виду, что формула (3.18) имеет место только для малых p_α, т. е. в случае $p_\alpha \ll \hbar/a$ (a — постоянная решетки), что обычно хорошо выполняется в полупроводниках из-за малой концентрации электронов проводимости. Вблизи точек $\boldsymbol{k}_0$, в которых $\varepsilon_n(\boldsymbol{k}_0)$ принимает экстремальное значение, изоэнергетические поверхности $\varepsilon_n(\boldsymbol{k}) = \text{const}$ имеют вид эллипсоидов.

Наконец, в том случае, когда все три компоненты тензора $1/m_\alpha$ одинаковы, можно ввести скалярную эффективную массу m^*, положив

$$\frac{1}{m^*} = \frac{1}{m_\alpha} = \frac{1}{\hbar^2}\left(\frac{\partial^2 \varepsilon_n}{\partial k_x^2}\right)_{\boldsymbol{k}_0} = \frac{1}{\hbar^2}\left(\frac{\partial^2 \varepsilon_n}{\partial k_y^2}\right)_{\boldsymbol{k}_0} = \frac{1}{\hbar^2}\left(\frac{\partial^2 \varepsilon_n}{\partial k_z^2}\right)_{\boldsymbol{k}_0}. \tag{3.22}$$

Энергия электрона в этом случае равна

$$\varepsilon_n(\boldsymbol{k}) = \hbar^2 k^2/2m^* = p^2/2m^*, \tag{3.22а}$$

т. е. равна кинетической энергии свободного электрона с массой m^* и импульсом $\boldsymbol{p}$.

У нижнего края зоны, там где $\varepsilon_n(\boldsymbol{k})$ достигает минимума, $\frac{1}{m_\alpha} = \frac{1}{\hbar^2}\left(\frac{\partial^2 \varepsilon_n}{\partial k_\alpha^2}\right)_{\boldsymbol{k}_0} > 0$, т. е. эффективная масса m_α положительна. Наоборот, вблизи максимума $\varepsilon_n(\boldsymbol{k})$ вторые производные $\left(\frac{\partial^2 \varepsilon_n}{\partial k_\alpha^2}\right)_{\boldsymbol{k}_0} < 0$ и эффективные массы отрицательны. Ниже мы увидим, что эффективные массы, так же как и обычные массы, определяют отношение силы к ускорению. Если $m_\alpha < 0$, то направление ускорения электрона противоположно направлению действующей на него силы. Это не должно удивлять нас, так как метод эффективной массы простейшим образом учитывает влияние на электрон периодического поля кристалла, а совместное действие периодического поля и внешней силы может для электрона, обладающего волновыми свойствами, приводить к указанному выше эффекту.

4. Если подставить выражение для энергии (3.21) в (3.17), то получим уравнение

$$-\left[\frac{\hbar^2}{2m_x}\frac{\partial^2}{\partial x^2} + \frac{\hbar^2}{2m_y}\frac{\partial^2}{\partial y^2} + \frac{\hbar^2}{2m_z}\frac{\partial^2}{\partial z^2} + \mathcal{U}(\boldsymbol{r})\right]\psi(\boldsymbol{r}) = \varepsilon\psi(\boldsymbol{r}), \tag{3.23}$$

описывающее движение электрона с неравными по осям массами m_x, m_y, m_z [1]) в поле $\mathcal{U}(\boldsymbol{r})$. Особенно простой вид имеет уравнение (3.23), когда тензор m_α^{-1} вырождается в скаляр

[1]) Необходимо помнить, что оси x, y и z совпадают с главными осями тензора $m_{\alpha\beta}^{-1}$, т. е. определенным образом ориентированы относительно кристалла.

$m_x = m_y = m_z = m^*$ и поле $\mathcal{U}(\boldsymbol{r}) = 0$, тогда

$$-\frac{\hbar^2}{2m^*}\nabla^2\psi(\boldsymbol{r}) = \varepsilon\psi(\boldsymbol{r}), \tag{3.23а}$$

что соответствует свободной частице с перенормированной массой ($m \to m^*$).

Такая ситуация может иметь место, когда минимум энергии $\varepsilon_n(\boldsymbol{k})$ находится в центре бриллюэновской зоны кубического кристалла.

Следует, однако, иметь в виду, что уравнение (3.23) было получено в предположении, что $\varepsilon_n(\boldsymbol{k})$ может быть заменено на (3.21) во всей бриллюэновской зоне, что, вообще говоря, неверно, даже в том случае, когда электроны движутся вблизи экстремума энергии.

Строгая теория, развитая В. Коном и И. Латтинджером (1955), обобщенная и на случай вырожденных зон, показывает, что уравнение (3.17) описывает только плавно изменяющуюся часть $F(\boldsymbol{r})$ волновой функции, а полная волновая функция электрона

$$\psi(\boldsymbol{r}) = F(\boldsymbol{r})\, u_{n\boldsymbol{k}_0}(\boldsymbol{r})\, e^{i\boldsymbol{k}_0\boldsymbol{r}}, \tag{3.24}$$

где $u_{n\boldsymbol{k}_0}(\boldsymbol{r})$ — амплитудный множитель функции Блоха (3.1) в точке $\boldsymbol{k} = \boldsymbol{k}_0$. С другой стороны, в некоторых случаях в качестве волновой функции электрона достаточно использовать ее плавно изменяющуюся часть $F(\boldsymbol{r})$, а собственные значения энергии электрона определять из уравнения (3.23).

5. В гл. II, § 8, п. 3 было показано, что энергетическая поверхность электрона в поле кристалла $\varepsilon_n(k) = \text{const}$, обладает симметрией точечной группы решетки $\mathcal{F}$; было также отмечено, что во всех случаях $\varepsilon_n(-\boldsymbol{k}) = \varepsilon_n(\boldsymbol{k})$, т. е. поверхность $\varepsilon_n(\boldsymbol{k}) = \text{const}$ обладает центром симметрии (независимо от того, обладает ли им кристаллическая решетка).

Таким образом, если, например, в кубической решетке в какой-либо произвольной точке $\boldsymbol{k}_0$ обратного пространства, не расположенной на каком-либо элементе симметрии, имеется экстремум $\varepsilon_n(\boldsymbol{k}_0)$, то в зоне Бриллюэна должно быть еще 48 симметрично расположенных экстремумов. Если $\boldsymbol{k}_0$ совпадает с направлением [100], то будет всего шесть экстремумов, а изоэнергетические поверхности будут эллипсоидами вращения (3.15) с осями симметрии, совпадающими с направлением [100]. Если центры этих эллипсоидов, т. е., другими словами, минимумы энергии ε_n, лежат в центрах граней зоны Бриллюэна, то половина каждого эллипсоида расположена за пределами первой зоны; на последнюю приходится шесть полуэллипсоидов, что эквивалентно трем целым эллипсоидам.

Аналогичная ситуация изображена для плоской квадратной решетки на рис. IV.1. На рис. IV.1, *а* и *б* показаны в $\boldsymbol{k}$-пространстве «эллипсы постоянной энергии» для случаев, когда минимумы энергии расположены в четырех эквивалентных (симметричных) точках внутри зоны Бриллюэна и на ее гранях. Рис. IV.1, *в* изображает периодическую повторяемость структуры, которая показана на рис. IV.1, *б*. Подобные и более сложные структуры энергетических зон являются не только теоретически возможными, но и действительно реализуются во многих веществах (кремний, германий и др.). Ниже мы остановимся на этом подробнее.

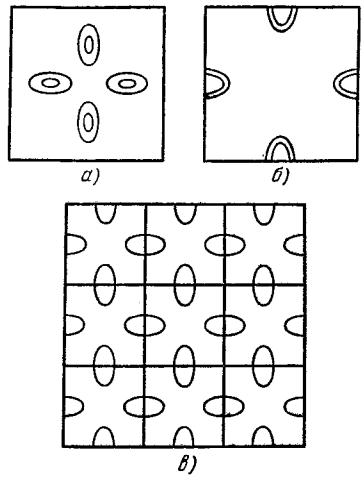


Рис. IV.1.

6. Условия цикличности Борна — Кармана, налагаемые на блоховскую волновую функцию электрона или на колебания атомов решетки, приводят к тому, что волновые векторы $\boldsymbol{k}$ и $\boldsymbol{q}$ принимают G^3 квазидискретных значений (гл. II, § 9, п. 1). Таким образом, можно говорить о числе состояний электрона в объеме V в области $\boldsymbol{k}$-пространства $d\tau_{\boldsymbol{k}}$ или в интервале энергии $(\varepsilon,\ \varepsilon+d\varepsilon)$.

Поскольку выражение для $\boldsymbol{q}$ (III.5.24) совпадает с выражением для $\boldsymbol{k}$ (3.2), число состояний электрона в интервале энергии $(\varepsilon,\ \varepsilon+d\varepsilon)$, если учесть (III.5.27), равно

$$g(\varepsilon)\,d\varepsilon=\frac{V}{(2\pi)^3}\sum_n \iiint\limits_{\varepsilon\leqslant\varepsilon_n(\boldsymbol{k})\leqslant\varepsilon+d\varepsilon} d\tau_{\boldsymbol{k}}, \tag{3.25}$$

а плотность состояний, аналогично (III.5.32), равна

$$g(\varepsilon)=\frac{V}{(2\pi)^3}\sum_n \int \frac{d\sigma}{|\nabla_{\boldsymbol{k}}\varepsilon_n|}, \tag{3.26}$$

где $\nabla_{\boldsymbol{k}}\varepsilon_n \equiv \operatorname{grad}_{\boldsymbol{k}}\varepsilon_n(\boldsymbol{k})$ и интегрирование ведется по поверхности $\varepsilon_n(\boldsymbol{k})=\text{const}$; вывод выражений (3.25) и (3.26) ничем не отличается от вычисления функции распределения частот в гл. III, § 5, п. 6; необходимо только заменить $\boldsymbol{q}\longrightarrow\boldsymbol{k}$ и $\omega_j\longrightarrow\varepsilon_n$.

Применим эту формулу к простому случаю одной зоны, когда энергия определяется выражением (3.22а) и, следовательно, изо-

энергетические поверхности — сферы. В этом случае

$$|\operatorname{grad}_{\boldsymbol{k}} \varepsilon| = \frac{d\varepsilon}{dk} = \frac{\hbar^2 k}{m^*},$$

а поверхность сферы $\sigma = 4\pi k^2$.

Плотность состояний в 1 $см^3$ ($V = 1\ см^3$) будет равна

$$g(\varepsilon) = \frac{1}{8\pi^3} \frac{m^*}{\hbar^2 k} 4\pi k^2 = \frac{\sqrt{2}}{2\pi^2} \frac{m^{*3/2}}{\hbar^3} \sqrt{\varepsilon} \sim \varepsilon^{1/2}, \qquad (3.27)$$

где мы опять воспользовались выражением (3.22а) для того, чтобы от $\boldsymbol{k}$ перейти к ε.

Этот же результат можно получить прямо из Приложения 4. Рассматривая формулу (П.4.1) для переменного импульса p и, понимая под $\Delta N = n_s$ число квантовых состояний для свободных электронов с импульсом, меньшим p, получим ($\Delta V = 1\ см^3$)

$$dn_s = \frac{1}{\pi^2} \frac{p^2 dp}{\hbar^3}.$$

Переходя к переменной $\varepsilon = p^2/2m^*$, получим для плотности состояний $\frac{dn_s}{d\varepsilon} = 2g(\varepsilon)$ формулу (3.27).

7. Вычислим квантовомеханическую среднюю скорость $\langle \boldsymbol{v} \rangle = \frac{1}{m} \langle \boldsymbol{p} \rangle$ электрона в состоянии с волновым вектором $\boldsymbol{k}$ в периодическом поле. Как мы увидим ниже, она в общем случае отлична от нуля. Таким образом, даже в том случае, когда полная энергия электрона меньше максимума его потенциальной энергии в кристалле, электрон может свободно перемещаться по всему объему кристалла. Электрону в состоянии $\boldsymbol{k}$ соответствует средняя скорость $\boldsymbol{v}(\boldsymbol{k})$ и, следовательно, незатухающий ток $\boldsymbol{j} = e\boldsymbol{v}$, где e — величина заряда электрона. Конечное электрическое сопротивление кристалла обусловливается не потенциальными барьерами периодического поля, которые электрон проходит «туннельным» способом, а отступлениями поля кристалла от строгой периодичности либо за счет тепловых колебаний, либо из-за статических дефектов решетки.

По общей формуле для средних значений в квантовой механике [1])

$$\langle \boldsymbol{v} \rangle = \frac{1}{m} \langle \boldsymbol{p} \rangle = \frac{1}{m} \frac{\hbar}{i} \int_V \psi_{\boldsymbol{k}}^* \nabla \psi_{\boldsymbol{k}} \, d\tau, \qquad (3.28)$$

где $\frac{\hbar}{i} \nabla \equiv \frac{\hbar}{i} \operatorname{grad}$ — оператор импульса; интегрирование ведется по основной области кристалла.

---

[1]) Блохинцев Д. И. Основы квантовой механики.— М.: Наука, 1976, § 19.

Для свободного электрона волновая функция, нормированная на основную область, равна

$$\psi_{\boldsymbol{k}}^{(0)} = \frac{1}{\sqrt{V}} e^{i\boldsymbol{k}\boldsymbol{r}}, \tag{3.29}$$

т. е. имеет вид плоской волны с постоянной амплитудой. Здесь волновой вектор $\boldsymbol{k} = \boldsymbol{p}/\hbar$, где $\boldsymbol{p}$ — импульс электрона (не квазиимпульс!). Подставляя (3.29) в (3.28), учитывая, что $\nabla e^{i\boldsymbol{k}\boldsymbol{r}} = i\boldsymbol{k}e^{i\boldsymbol{k}\boldsymbol{r}}$, получим (опуская скобки у $\langle \boldsymbol{v} \rangle$)

$$\boldsymbol{v} = \frac{\hbar}{mi} \int \psi_{\boldsymbol{k}}^{(0)*} i\boldsymbol{k} \psi_{\boldsymbol{k}}^{(0)} \, d\tau = \frac{\hbar \boldsymbol{k}}{m} = \frac{\boldsymbol{p}}{m}, \tag{3.30}$$

т. е. то же соотношение, что и в классической механике.

Если отождествить (что à priori не очевидно) среднюю квантовомеханическую скорость электрона в кристалле $\boldsymbol{v}$ с групповой скоростью волнового пакета, составленного из блоховских функций $\boldsymbol{v}_{\text{гр}}$, то [1])

$$\boldsymbol{v}_{\text{гр}} = \operatorname{grad}_{\boldsymbol{k}} \omega = \frac{\partial \omega}{\partial \boldsymbol{k}} = \frac{1}{\hbar} \frac{\partial \hbar\omega}{\partial \boldsymbol{k}} = \frac{1}{\hbar} \frac{\partial \varepsilon}{\partial \boldsymbol{k}} \tag{3.31}$$

и, следовательно,

$$\boldsymbol{v} = \frac{1}{\hbar} \frac{\partial \varepsilon}{\partial \boldsymbol{k}} = \frac{1}{\hbar} \operatorname{grad}_{\boldsymbol{k}} \varepsilon(\boldsymbol{k}). \tag{3.32}$$

В случае свободного электрона энергия

$$\varepsilon(\boldsymbol{k}) = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m} \tag{3.33}$$

и формула (3.32) дает (3.30).

В точках экстремума энергии $(\partial\varepsilon/\partial \boldsymbol{k}) = 0$ и, следовательно, $\boldsymbol{v} = 0$.

Строгий вывод формулы (3.32) дан в Приложении 12.

8. Рассмотрим электрон в периодическом поле, когда на него дополнительно действует внешняя сила $\boldsymbol{F}$. Если сила $\boldsymbol{F}$ достаточно мала, так что $Fa \ll \mathscr{E}_g$, где a — постоянная решетки, а $\mathscr{E}_g$ — ширина соответствующей запрещенной зоны, то сила F не будет вызывать переходов электрона между разными энергетическими зонами, а будет только изменять волновой вектор электрона $\boldsymbol{k}$. Так как классические уравнения движения остаются справедливыми для средних квантовомеханических значений величин, то естественно предположить, что закон сохранения энергии имеет вид

$$\frac{d\varepsilon(\boldsymbol{k})}{dt} = \boldsymbol{v}\boldsymbol{F}, \tag{3.34}$$

где $\boldsymbol{v}$ — скорость, определяемая равенством (3.32), а $\varepsilon(\boldsymbol{k})$ — энергия электрона в зоне, в которой он движется. Равенство (3.34)

[1]) Блохинцев Д. И., § 7.

утверждает, что работа силы $\boldsymbol{F}$, приложенной к электрону, в 1 *сек* равна скорости изменения его энергии $\varepsilon(\boldsymbol{k})$.

Так как

$$\frac{d\varepsilon(\boldsymbol{k})}{dt}=\sum_{\alpha}\frac{\partial\varepsilon}{\partial k_\alpha}\frac{dk_\alpha}{dt}=\operatorname{grad}_{\boldsymbol{k}}\varepsilon\frac{d\boldsymbol{k}}{dt},$$

то, используя равенство (3.32), получим из уравнения (3.34)

$$\frac{d(\hbar\boldsymbol{k})}{dt}=\frac{d\boldsymbol{p}}{dt}=\boldsymbol{F}. \tag{3.35}$$

Отсюда, в частности, видно, что в периодическом поле квазиимпульс электрона $\boldsymbol{p}=\hbar\boldsymbol{k}$ играет в уравнении движения (3.35) роль импульса свободного электрона.

Ускорение электрона, понимаемое как скорость изменения его средней квантовомеханической скорости $\boldsymbol{v}$, равно

$$\frac{d\boldsymbol{v}}{dt}=\frac{d}{dt}\left(\frac{1}{\hbar}\operatorname{grad}_{\boldsymbol{k}}\varepsilon(\boldsymbol{k})\right)=\frac{1}{\hbar}\frac{d}{dt}\left(\frac{\partial\varepsilon(\boldsymbol{k})}{\partial\boldsymbol{k}}\right).$$

Так как $\frac{\partial\varepsilon}{\partial\boldsymbol{k}}$ зависит от времени только через $\boldsymbol{k}$, то

$$\frac{dv_\alpha}{dt}=\frac{1}{\hbar}\frac{d}{dt}\left(\frac{\partial\varepsilon}{\partial k_\alpha}\right)=\frac{1}{\hbar}\sum_{\beta}\left(\frac{\partial^2\varepsilon}{\partial k_\alpha\partial k_\beta}\right)\frac{dk_\beta}{dt}=\sum_{\beta}\frac{1}{\hbar^2}\left(\frac{\partial\varepsilon}{\partial k_\alpha\partial k_\beta}\right)\frac{dp_\beta}{dt}. \tag{3.36}$$

Совокупность величин

$$\frac{1}{\hbar^2}\left(\frac{\partial^2\varepsilon}{\partial k_\alpha\partial k_\beta}\right)=\tilde{m}^{-1}_{\alpha\beta} \tag{3.37}$$

назовем *обобщенным тензором обратной эффективной массы*. Он отличается от тензора $m^{-1}_{\alpha\beta}$ (3.19) тем, что зависит от $\boldsymbol{k}$. Если энергия электрона равна (3.18), т. е. берется в квадратичном приближении, то $\tilde{m}^{-1}_{\alpha\beta}=m^{-1}_{\alpha\beta}$. Из уравнений (3.35) и (3.36) в главных осях тензора $\tilde{m}^{-1}_{\alpha\beta}$ следует:

$$\frac{dv_\alpha}{dt}=\tilde{m}^{-1}_{\alpha}F_\alpha, \tag{3.38}$$

что в случае $\tilde{m}^{-1}_{\alpha\beta}=m^{-1}_{\alpha}=\text{const}$ представляет классические уравнения движения с «анизотропной» массой.

Наконец, в приближении скалярной эффективной массы (3.22) уравнение (3.38) имеет вид

$$m^*\frac{d\boldsymbol{v}}{dt}=\boldsymbol{F}, \tag{3.39}$$

совпадающий с обычным уравнением классической механики.

§ 4. Понятие о положительных дырках почти заполненной валентной зоны

1. По аналогии с классическим выражением для тока, создаваемого зарядом $-e$, движущимся со скоростью $\boldsymbol{v}$, напишем квантовомеханическое выражения для тока, создаваемого электроном в $\boldsymbol{k}$-состоянии:

$$\boldsymbol{j}_{\boldsymbol{k}} = -e\boldsymbol{v}(\boldsymbol{k}) = -\frac{e\hbar}{mi}\int \psi_{\boldsymbol{k}}^{*}\nabla\psi_{\boldsymbol{k}}\,d\tau = -\frac{e\hbar}{2mi}\int (\psi_{\boldsymbol{k}}^{*}\nabla\psi_{\boldsymbol{k}} - \psi_{\boldsymbol{k}}\nabla\psi_{\boldsymbol{k}}^{*})\,d\tau. \tag{4.1}$$

Последний интеграл может быть получен, если учесть, что формально $\boldsymbol{j}_{\boldsymbol{k}} + \boldsymbol{j}_{\boldsymbol{k}}^{*} = 2\boldsymbol{j}_{\boldsymbol{k}}$ ($\boldsymbol{j}_{\boldsymbol{k}}$ — вещественная величина). Мы видим, что правая часть выражения (4.1) действительно представляет собой обычное квантовомеханическое выражение для плотности тока, усредненное по основной области кристалла (строго говоря, для усреднения интеграл в (4.1) надо было бы разделить на V). Такое усреднение устраняет циркулярные токи внутри отдельных кристаллических ячеек, токи, которые связаны с наличием периодического множителя $u_{\boldsymbol{k}}(\boldsymbol{r})$ в волновой функции Блоха, не имеющие отношения к поступательному движению электрона [1]).

Мы знаем, что в кристалле любой симметрии (П.9.26)

$$\varepsilon(-\boldsymbol{k}) = \varepsilon(\boldsymbol{k}) \tag{4.2}$$

и, следовательно,

$$\boldsymbol{v}(-\boldsymbol{k}) = -\boldsymbol{v}(\boldsymbol{k}). \tag{4.3}$$

Последнее непосредственно вытекает из выражений для средней скорости (3.32).

Используя равенство (4.3), легко показать, что

$$\sum_{\boldsymbol{k}} \boldsymbol{v}(\boldsymbol{k}) = 0, \tag{4.4}$$

где суммирование распространяется на все значения $\boldsymbol{k}$ внутри зоны Бриллюэна. Согласно принципу Паули в каждом $\boldsymbol{k}$-состоянии может находиться не больше двух электронов с противоположно направленными спинами s_z. Таким образом, ток полностью заполненной валентной зоны

$$\boldsymbol{j} = \sum_{\boldsymbol{k}s} \boldsymbol{j}_{\boldsymbol{k}} = 2\sum_{\boldsymbol{k}} \boldsymbol{j}_{\boldsymbol{k}} = -e2\sum_{\boldsymbol{k}} v(\boldsymbol{k}) = 0. \tag{4.5}$$

Если к кристаллу приложить электрическое поле $\boldsymbol{E} = -\operatorname{grad}\varphi$, то ток полностью заполненной зоны по-прежнему равен нулю. В самом деле, если поле не слишком велико, то электроны не будут перебрасываться в зону проводимости и валентная зона

[1]) См. Spenke E. Elektronische Halbleiter.— Berlin: Springer-Verlag, 1956, p. 196.

останется заполненной и в электрическом поле. Если зона заполнена электронами не полностью, но внешнее электрическое поле равно нулю, то ток

$$\boldsymbol{j}=\sum_{ks}{}' \boldsymbol{j}_k=0 \tag{4.6}$$

из-за симметрии распределения электронов в $\boldsymbol{k}$-пространстве (штрих у суммы означает, что суммирование ведется не по всей зоне Бриллюэна). Но если электроны не полностью заполняют зону и к кристаллу приложено электрическое поле, то ток (4.6) отличен от нуля, так как число $\boldsymbol{k}$-состояний, заполненных электронами, движущимися против поля, возрастает.

Введем теперь символ $\nu_n(\boldsymbol{k}, s)$, равный единице, если состояние $(\boldsymbol{k}, s)$ занято электроном, и нулю — в противном случае. Очевидно, вероятность того, что состояние $(\boldsymbol{k}, s)$ не занято электроном (дырка), равна $\nu_p(\boldsymbol{k}, s)=1-\nu_n(\boldsymbol{k}, s)$. Ток (4.6) может быть теперь записан так:

$$\boldsymbol{j}=-e\sum_{ks}\nu_n(\boldsymbol{k}, s)\,\boldsymbol{v}(\boldsymbol{k}), \tag{4.6а}$$

где суммирование ведется по всей зоне, или

$$\boldsymbol{j}=-e\sum_{ks}[1-\nu_p(\boldsymbol{k}, s)]\,\boldsymbol{v}(\boldsymbol{k})=-e\sum_{ks}\boldsymbol{v}(\boldsymbol{k})+\sum_{ks}\nu_p(\boldsymbol{k}, s)\,\boldsymbol{v}(\boldsymbol{k})=$$
$$=e\sum_{ks}\nu_p(\boldsymbol{k}, s)\,\boldsymbol{v}(\boldsymbol{k}). \tag{4.6б}$$

Мы видим, что электрический ток неполностью заполненной электронами зоны можно описать как ток положительно заряженных $+e$ квазичастиц или дырок, соответствующих незанятым электронами $\boldsymbol{k}, s$-состояниям, движущихся со скоростями $\boldsymbol{v}(\boldsymbol{k})$. Концентрация дырок (их количество в 1 $см^3$) равна $\sum_{ks}\nu_p(\boldsymbol{k}, s)$, если число квантовых состояний берется в расчете на 1 $см^3$. Утверждение, что под действием электрического поля увеличивается число $\boldsymbol{k}$-состояний, занятых электронами, скорость которых $v(\boldsymbol{k})$ направлена п р о т и в поля, эквивалентно утверждению, что увеличивается число $\boldsymbol{k}$-состояний, з а н я т ы х д ы р к а м и, скорость которых направлена по полю. Вычислим поток энергии $\boldsymbol{w}$, переносимый электронами неполностью заполненной зоны, если имеется преимущественное направление в их движении (обусловленное градиентом температуры, концентрации или потенциала φ):

$$\boldsymbol{w}=\sum_{ks}\nu_n(\boldsymbol{k}, s)[\varepsilon(\boldsymbol{k})-e\varphi]\,\boldsymbol{v}(\boldsymbol{k})=\sum_{ks}[1-\nu_p(\boldsymbol{k}, s)][\varepsilon(\boldsymbol{k})-e\varphi]\,\boldsymbol{v}(\boldsymbol{k})=$$
$$=\sum_{ks}[\varepsilon(\boldsymbol{k})-e\varphi]\boldsymbol{v}(\boldsymbol{k})+\sum_{ks}\nu_p(\boldsymbol{k}, s)[-\varepsilon(\boldsymbol{k})+e\varphi]\,\boldsymbol{v}(\boldsymbol{k}). \tag{4.7}$$

Первая сумма в силу соотношений (4.2) и (4.3) равна нулю, вторая сумма может быть интерпретирована как поток энергии,

переносимый дырками с зарядом $+e$ и энергией $-\varepsilon(\boldsymbol{k})$. Если зона почти заполнена, так что энергия дырок при $\boldsymbol{k} \approx \boldsymbol{k}_0$ близка к верхнему краю зоны, то в приближении эффективной массы (3.21) имеем

$$-\varepsilon(\boldsymbol{k}) = -\varepsilon(\boldsymbol{k}_0) + \sum_{\alpha} \frac{\hbar^2 k_\alpha^2}{2(-m_\alpha)}, \qquad (4.7а)$$

где эффективная масса дырки $-m_\alpha$ положительна, так как эффективная масса электрона m_α у верхнего края зоны отрицательна; таким образом,

$$m_\alpha^{(p)} = -m_\alpha^{(n)} > 0. \qquad (4.7б)$$

При наличии магнитного поля $\boldsymbol{H}$ на электрон действует сила Лоренца $-\frac{e}{c}[\boldsymbol{v}\boldsymbol{H}]$, так же как и в случае электрического поля пропорциональная заряду $-e$; поэтому можно показать, что и в этом случае остается применимой концепция дырок. В дальнейшем будет показано, что статистическое поведение электронов почти заполненной зоны эквивалентно статистическому поведению дырок, уровни энергии которых определяются выражением (4.7а). Таким образом, как в оптических явлениях, так и в явлениях переноса, обусловленных электрическими и магнитными полями или градиентом температуры или концентрации, можно пользоваться представлением о дырках.

2. В полупроводниках в общем случае носителями тока являются как электроны нижней части зоны проводимости, так и электроны почти заполненной валентной зоны, которые можно интерпретировать как квазичастицы с положительным зарядом $+e$ и положительными массами (4.7б). Часто в определенном температурном интервале можно считать, что явления переноса осуществляются либо одними электронами зоны проводимости (электронный или n-полупроводник), либо дырками валентной зоны (дырочной или p-полупроводник). Очевидно, что изучение электропроводности полупроводника не позволяет нам определить, принадлежит ли он к n- или p-типу. Как мы увидим дальше, это можно решить, изучая холл-эффект и другие гальвано- и термомагнитные явления.

Следует заметить, что концепция электронов и дырок с некоторыми эффективными массами может быть обоснована только в том случае, когда на электроны проводимости в кристалле действует внешнее электромагнитное поле[1]). Если же на электроны проводимости в кристалле действует гравитационное поле или силы инерции, то концепция дырок и электронов с эффективными массами неприменима. Наиболее прямым опытом для

[1]) Darwin G. G.— Proc. Roy. Soc.— London, 1936, v. A 154, p. 61.

обнаружения сил инерции, действующих на электроны проводимости, является опыт Стюарта и Толмена (1916). Идея опыта состоит в следующем: цилиндрическая катушка с большим числом витков приводится в быстрое вращение вокруг своей оси. При резком торможении катушки электроны проводимости продолжают двигаться по инерции и в ней возникает кратковременный ток.

Нетрудно показать[1]), что полный заряд q, протекающий в цепи за все время торможения катушки, равен

$$q = \frac{m}{e}\frac{v_0 l}{R}, \tag{4.8}$$

где v_0 — начальная линейная скорость провода, l — его длина, R — сопротивление цепи, e и m — заряд и масса частиц, на которые действуют силы инерции. Измеряя заряд q баллистическим гальванометром, легко определить отношение e/m (при заданных l, R и v_0).

Опыт, в согласии с теорией, показал, что отношение e/m по знаку и величине всегда совпадает со значением для свободных электронов. Поэтому попытки обнаружить в опытах типа Стюарта—Толмена движение дырок с эффективными массами m^*, были заранее обречены на неудачу[2]).

3. Приведенное выше рассмотрение дырок может привести к неправильному представлению, что концепция дырок полностью эквивалентна незаполненной электронами зоне независимо от степени ее заполнения. Следует, однако, обратить внимание на то, что переход к дыркам происходит путем формальной замены под знаком суммы $\sum_{ks}$ символа $\nu_n(\boldsymbol{k}, s)$ на $1-\nu_p(\boldsymbol{k}, s)$, т. е. физически путем замены совокупности электронов, частично заполняющих зону, на совокупность электронов, полностью заполняющих зону, плюс дырки.

В балансе явлений переноса обе эти совокупности действительно эквивалентны, но нет оснований считать их эквивалентными с точки зрения самосогласованного периодического потенциала одноэлектронной задачи.

В последнем случае эквивалентность будет выполняться тем лучше, чем меньше концентрация дырок, т. е. чем больше заполнена электронами валентная зона.

---

[1]) С. Г. Калашников. Электричество.— 4 изд.— М., Наука, 1977, § 145.

[2]) Brown S., Barnett S. J.— Phys. Rev., 1952, v. 87, p. 601.

§ 5. Приближение почти свободных (слабо связанных) электронов

Рассмотрим движение электрона в слабом периодическом поле $V(\boldsymbol{r})$, т. е. будем считать $V(\boldsymbol{r})$ малым возмущением свободного движения электрона (Пайерлс) [1]).

Разложим периодический потенциал в ряд Фурье (I.3.6):

$$V(\boldsymbol{r}) = \sum_{\boldsymbol{g} \neq 0} V_{\boldsymbol{g}} e^{i(\boldsymbol{b}_{\boldsymbol{g}} \boldsymbol{r})}, \tag{5.1}$$

где $\boldsymbol{b}_{\boldsymbol{g}}$ — вектор обратной решетки. Не ограничивая общности, можно положить нулевой член в разложении (5.1), т. е. среднее значение потенциала $V_0 = 0$. Для вещественности правой части (5.1) необходимо, чтобы $V_{-\boldsymbol{g}} = V_{\boldsymbol{g}}^*$. Будем считать амплитуды $V_{\boldsymbol{g}}$ величинами первого порядка малости. Разложим в ряд Фурье периодический множитель блоховской волновой функции:

$$u_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{h}} a_{\boldsymbol{h}} e^{i(\boldsymbol{b}_{\boldsymbol{h}} \boldsymbol{r})}, \tag{5.2}$$

где $\boldsymbol{b}_{\boldsymbol{h}}$ — вектор обратной решетки.

Подставим выражения (5.1) и (5.2) в уравнение (3.8) для $u_{\boldsymbol{k}}(\boldsymbol{r})$; получим

$$\sum_{\boldsymbol{h}} \frac{\hbar^2}{2m} \boldsymbol{b}_{\boldsymbol{h}}^2 a_{\boldsymbol{h}} e^{i(\boldsymbol{b}_{\boldsymbol{h}} \boldsymbol{r})} + \sum_{\boldsymbol{h}} \sum_{\boldsymbol{g} \neq 0} V_{\boldsymbol{g}} a_{\boldsymbol{h}} e^{i(\boldsymbol{b}_{\boldsymbol{g}+\boldsymbol{h}}, \boldsymbol{r})} + \\ + \sum_{\boldsymbol{h}} \frac{\hbar^2}{m} (\boldsymbol{b}_{\boldsymbol{h}} \boldsymbol{k}) a_{\boldsymbol{h}} e^{i(\boldsymbol{b}_{\boldsymbol{h}} \boldsymbol{r})} = \sum_{\boldsymbol{h}} \left(\varepsilon_{\boldsymbol{k}} - \frac{\hbar^2 \boldsymbol{k}^2}{2m} \right) a_{\boldsymbol{h}} e^{i(\boldsymbol{b}_{\boldsymbol{h}} \boldsymbol{r})}. \tag{5.3}$$

В двойной сумме суммирование по $\boldsymbol{h}$ и $\boldsymbol{g}$ заменим суммированием по $\boldsymbol{h} - \boldsymbol{g}$ и $\boldsymbol{g}$. В этом случае в экспоненте вектор $\boldsymbol{b}_{\boldsymbol{g}} + \boldsymbol{b}_{\boldsymbol{h}}$ надо заменить на $\boldsymbol{b}_{\boldsymbol{h}}$, так что двойная сумма примет вид

$$\sum_{\boldsymbol{h}} \sum_{\boldsymbol{g}} V_{\boldsymbol{g}} a_{\boldsymbol{h}-\boldsymbol{g}} e^{i(\boldsymbol{b}_{\boldsymbol{h}} \boldsymbol{r})}. \tag{5.3а}$$

Для того чтобы равенство (5.3) выполнялось тождественно для всех $\boldsymbol{r}$, необходимо, чтобы сумма коэффициентов при всех $\exp i(\boldsymbol{b}_{\boldsymbol{h}} \boldsymbol{r})$ равнялась нулю.

Учитывая (5.3а) и производя несложное алгебраическое преобразование, получим

$$\left[\varepsilon_{\boldsymbol{k}} - \frac{\hbar^2}{2m} (\boldsymbol{k} + \boldsymbol{b}_{\boldsymbol{h}})^2 \right] a_{\boldsymbol{h}} - \sum_{\boldsymbol{g} \neq 0} V_{\boldsymbol{g}} a_{\boldsymbol{h}-\boldsymbol{g}} = 0 \\ (h_1, h_2, h_3 = \pm 1, \pm 2, \pm 3, \ldots). \tag{5.4}$$

[1]) Другими словами, колебания потенциальной энергии электрона $V(\boldsymbol{r})$ малы по сравнению с его кинетической энергией.

Мы получили бесконечную линейную однородную систему алгебраических уравнений для неизвестных коэффициентов a_h. Обращаясь с системой (5.4) как с конечной, потребуем равенства нулю ее бесконечного определителя. Таким образом, мы получим уравнение, из которого в принципе можно определить спектр собственных значений энергии электрона ε_k, а затем из уравнений (5.4) — соответствующую систему коэффициентов a_h, т. е. волновую функцию электрона. Конечно, даже при известных коэффициентах V_g, т. е. при заданном периодическом потенциале, реальное получение решений безнадежно сложно.

Рассмотрим свободный электрон, когда все $V_g = 0$. В этом случае из уравнений (5.4) следует:

$$\varepsilon_k - \frac{\hbar^2}{2m}(\boldsymbol{k}+\boldsymbol{b}_h)^2 = 0, \text{ либо } a_h = 0. \tag{5.5}$$

Кроме того, для свободного электрона

$$\varepsilon_k = \varepsilon_0(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m. \tag{5.5а}$$

Из условий (5.5) и (5.5а) следует, что

$$a_0 \neq 0, \quad a_h = 0 \quad (\boldsymbol{h} \neq 0). \tag{5.5б}$$

Положим $a_0 = 1$, что соответствует нормировке волновой функции свободного электрона на единицу объема.

Решим теперь систему (5.4) в случае слабого периодического поля, когда коэффициенты V_g можно рассматривать как величины первого порядка малости. Естественно считать (что подтвердится расчетом), что в этом случае $a_{h\neq 0} \neq 0$ и тоже будут величинами первого порядка малости. Считая с точностью до величин первого порядка малости, удержим в левой части системы (5.4) одно слагаемое с $\boldsymbol{g} = \boldsymbol{h}$ (пропорциональное $a_0 = 1$) и подставим вместо ε_k его значение (5.5а); тогда

$$a_h = -\frac{2m}{\hbar^2}\frac{V_h}{\boldsymbol{b}_h^2 + 2(\boldsymbol{b}_h\boldsymbol{k})} \quad (\boldsymbol{h} \neq 0). \tag{5.6}$$

Совокупность коэффициентов $a_{h\neq 0}$ (5.6) определяет поправку к волновой функции электрона в первом приближении.

Для определения возмущенной энергии положим

$$\varepsilon_k = \varepsilon_0 + \varepsilon' = \frac{\hbar^2\boldsymbol{k}}{2m} + \varepsilon'. \tag{5.7}$$

Подставляя выражения (5.7) и (5.6) в уравнение (5.4), получим

$$\varepsilon' a_h - V_h = \sum_{g\neq 0} V_g a_{h-g}. \tag{5.8}$$

Как мы увидим, ε' — второго порядка малости; поэтому полученное равенство достаточно рассмотреть для $\boldsymbol{h} = 0$ (величины

$\varepsilon' a_{h \neq 0}$ — третьего порядка малости). В этом случае

$$\varepsilon' = \sum_{g \neq 0} V_g a_{-g}. \tag{5.9}$$

Заменяя суммирование по g суммированием по $-g$, подставляя a_g и учитывая, что $V_{-g} = V_g^*$, получим

$$\varepsilon' = -\frac{2m}{\hbar^2} \sum_{g \neq 0} \frac{|V_g|^2}{b_g^2 + 2(b_g k)}. \tag{5.9a}$$

Мы видим, что поправка к энергии ε' пропорциональна $|V_g|^2$, т. е. второго порядка малости, как мы утверждали выше.

Рассмотрим важный случай, когда знаменатели в выражениях (5.6) и (5.9а) близки к нулю, т. е.

$$^1/_2 b_g^2 + (b_g k) = (b_g,\ ^1/_2 b_g + k) \approx 0. \tag{5.10}$$

Очевидно, что соответствующий коэффициент a_g и поправка к энергии ε' не будут уже в этом случае малыми, и, следовательно, сами формулы (5.6) и (5.9а) неприменимы. Надо думать, что если волновой вектор электрона k удовлетворяет для некоторого вектора b_g условию (5.10), то почти свободное движение электрона претерпевает сильное возмущение. Это возмущение имеет характер зеркального отражения электронной волны по Вульфу—Брэггу от атомной плоскости, перпендикулярной к вектору b_g, так как условие (5.10) в точности совпадает с (I.4.5). В этом явлении особенно наглядно проявляется волновая природа электрона, движущегося в периодическом поле кристалла.

С другой стороны, условие (5.10) совпадает с уравнением плоскости, ограничивающей зоны Бриллюэна (II.9.11). Таким образом, если конец волнового вектора электрона k лежит вблизи границы зоны Бриллюэна, то электрон испытывает в кристалле сильное брэгговское отражение.

Рассмотрим электрон с волновым вектором k, удовлетворяющим условию (5.10), для заданного b_g. Если соответствующая амплитуда V_g в уравнении (5.6) не близка к нулю, то коэффициент a_g велик (одновременно с коэффициентом a_0). Система уравнений (5.4) сводится в этом случае к двум уравнениям:

$$\begin{aligned} \left[\varepsilon_k - \frac{\hbar^2 k^2}{2m}\right] a_0 &= V_{-g} a_g, \\ \left[\varepsilon_k - \frac{\hbar^2}{2m}(k + b_g)^2\right] a_g &= V_g a_0. \end{aligned} \tag{5.11}$$

Система линейных однородных алгебраических уравнений (5.11) для коэффициентов a_0 и a_g имеет решения, отличные от

нулевого, если определитель системы равен нулю, т. е.

$$\left[\varepsilon_{\boldsymbol{k}}-\frac{\hbar^2\boldsymbol{k}^2}{2m}\right]\left[\varepsilon_{\boldsymbol{k}}-\frac{\hbar^2}{2m}(\boldsymbol{k}+\boldsymbol{b}_g)^2\right]-|V_g|^2=0. \qquad (5.11a)$$

Решая это квадратное уравнение для $\varepsilon_{\boldsymbol{k}}$ и подставляя в решение (5.10), получим

$$\varepsilon_{\boldsymbol{k}}=\frac{\hbar^2\boldsymbol{k}^2}{2m}\pm|V_g|. \qquad (5.12)$$

Таким образом, для $\boldsymbol{k}$, удовлетворяющих интерференционному условию (5.10), энергия терпит разрыв, равный $2|V_g|$.

Рассмотрим для иллюстрации одномерный случай, когда $b_g=2\pi g/a$ и условие (5.10) приобретает вид

$$\frac{1}{2}\frac{4\pi^2g^2}{a^2}\pm\frac{k2\pi}{a}=0.$$

Отсюда

$$k=k_g=\pm\frac{\pi}{a}g \quad (g=1, 2, 3, \ldots). \qquad (5.13)$$

На рис. IV.2 изображена зависимость энергии $\varepsilon_{\boldsymbol{k}}$ от k для этого случая. При k, далеких от k_g, поправка к энергии мала

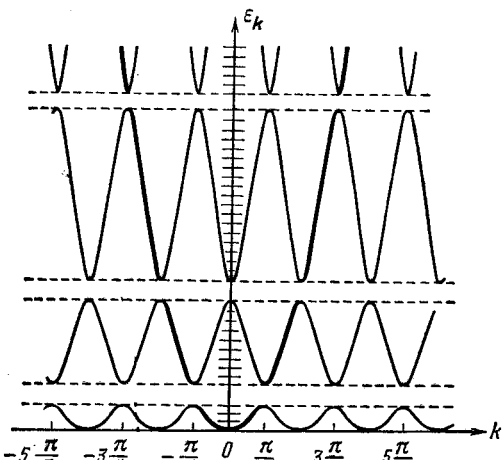


Рис. IV.2.

(квадратична в $|V_g|$), т. е. можно считать $\varepsilon_{\boldsymbol{k}}=\frac{\hbar^2k^2}{2m}$, где m — масса свободного электрона. При $k=k_g=\pm\pi/a$, $\pm 2\pi/a$, $\pm 3\pi/a$ в энергетическом спектре электрона возникают запрещенные участки энергии шириною $2|V_1|$, $2|V_2|$, $2|V_3|$ и т. д. Спектр энергии электрона приобретает зонный характер, т. е. разрешенные участки энергии чередуются с запрещенными. Мы видим,

что в случае почти свободных электронов их энергетический спектр имеет почти параболический характер (5.5а). В соответствии с условием (3.9) энергия электрона в пределах каждой разрешенной зоны — периодическая функция волнового вектора k, т. е. $\varepsilon_{k+\frac{2\pi}{a}} = \varepsilon_k$ (тонкие линии, продолжающие влево и вправо отрезки параболы на рис. IV.2). Это позволяет рассматривать все энергетические полосы в пределах первой или приведенной зоны Бриллюэна, т. е. для $-\frac{\pi}{a} \leqslant k \leqslant \frac{\pi}{a}$. Мы видим, что для слабо связанных электронов в приведенной зоне энергия электрона в последовательных энергетических зонах при $k=0$ попеременно принимает минимальное и максимальное значения. Одной из характерных особенностей одномерного случая, представленного на рис. IV.2, является то, что последовательные зоны разрешенной энергии электрона всегда разделены запрещенными участками энергии $2|V_1|$, $2|V_2|$ и т. д.

В двумерном и трехмерном случаях это не всегда имеет место. Теория и опыт показывают, что в этих случаях энергия электрона в некоторой точке бриллюэновской зоны для верхней полосы (зоны) энергии может оказаться ниже энергии, вообще говоря, в другой точке нижней полосы (зоны) энергии. В таком случае мы говорим о перекрытии зон энергии. Такое перекрытие энергетических зон играет важную роль в металлах. В этих случаях может начать заполняться электронами верхняя энергетическая зона при неполностью заполненной нижней зоне.

§ 6. Зоны Бриллюэна

1. В гл. II, § 9, п. 1 мы рассмотрели, как геометрически строятся зоны Бриллюэна. Для этого произвольный узел обратной решетки O соединяется отрезками с остальными узлами решетки и затем через середины этих отрезков проводятся перпендикулярные к ним плоскости. Многогранники, ограниченные этими плоскостями, образуют зоны Бриллюэна.

Для плоской квадратной решетки со стороною a обратная решетка тоже квадратная со стороною $2\pi/a$. На рис. IV.3 изображено 10 первых бриллюэновских зон для плоской квадратной решетки.

Для трехмерной простой кубической решетки обратная решетка — тоже простая кубическая. Первая зона Бриллюэна, имеющая форму куба, получается от пересечения шести плоскостей, проходящих через середины отрезков, соединяющих начало координат с шестью ближайшими узлами обратной кубической решетки. В огранении второй зоны участвуют двенадцать плоскостей, перпендикулярных к отрезкам, соединяющим начало

координат с двенадцатью следующими узлами обратной решетки. На рис. IV.4 представлены первые четыре бриллюэновские зоны простой кубической решетки.

В гл. II, § 9, п. 1 мы отметили, что все бриллюэновские зоны имеют одинаковый объем, равный $(2\pi)^3/\Omega_0$, где Ω_0 — объем

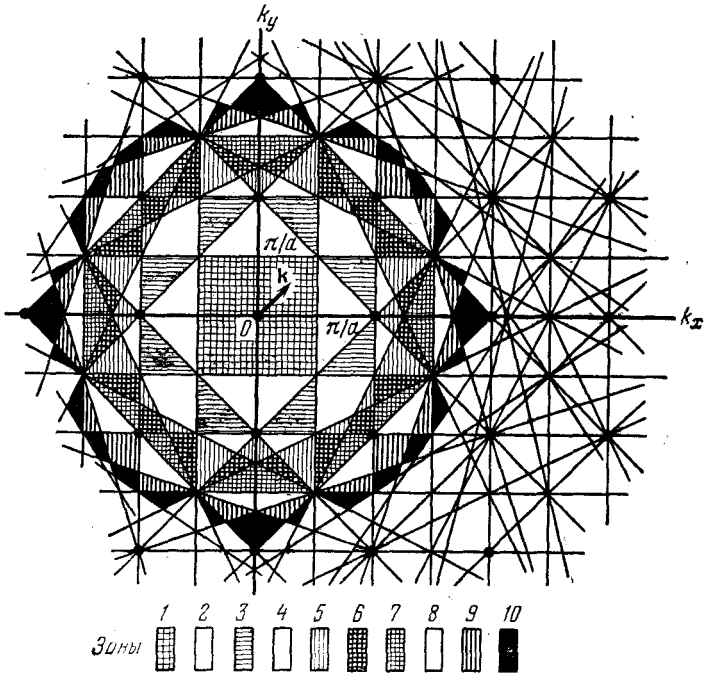


Рис. IV.3.

элементарной ячейки кристалла. С другой стороны, из (III.5.27) следует, что число состояний в основной области V на единицу объема пространства волнового вектора равно $V/(2\pi)^3$. Таким образом, число квантовых состояний электрона (без учета спина) во всех бриллюэновских зонах одинаково и равно $\frac{V}{(2\pi)^3}\frac{(2\pi)^3}{\Omega_0} = \frac{V}{\Omega_0} = G^3 = N$ — числу элементарных ячеек в основной области кристалла.

В качестве используемого в дальнейшем примера рассмотрим зоны Бриллюэна для гранецентрированной и объемноцентрированной кубической решеток.

Можно показать (Приложение 13), что для гранецентрированной кубической решетки обратная решетка — объемноцентри-

рованный куб и, наоборот, для прямой кубической объемноцентрированной решетки обратная — гранецентрированный куб. Определим форму 1-й зоны Бриллюэна для решетки гранецент-

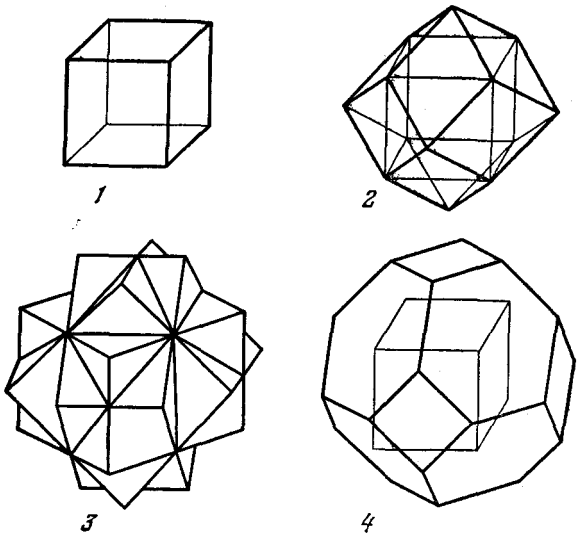


Рис. IV. 4.

рированного куба. Обратная решетка в этом случае объемноцентрированный куб, и, следовательно, каждый узел ее окружен восемью ближайшими узлами (гл. I, § 2, п. 2). Восемь плоскостей, проведенных перпендикулярно через середины отрез-

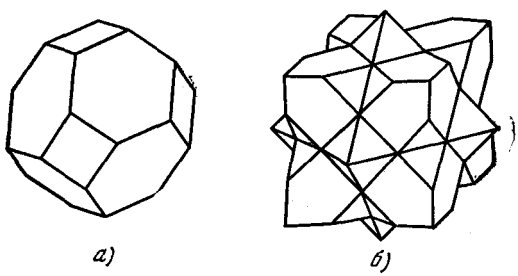


Рис. IV. 5.

ков, соединяющих начало координат с восемью ближайшими узлами, при их взаимном пересечении определяют правильный восьмигранник (октаэдр) с шестью вершинами и правильными треугольниками в виде граней. Узлами следующей координационной группы являются шесть узлов, расположенных попарно

на расстояниях a (ребро куба) по осям x, y и z. Они определяют шесть плоскостей, отсекающих от октаэдра его шесть вершин. В результате для 1-й зоны Бриллюэна мы получим четырнадцатигранник с шестью квадратными и восемью шестиугольными гранями, изображенный на рис. IV.5, *а*. Вторая зона Бриллюэна в этом случае имеет весьма сложную геометрическую форму, изображенную на рис. IV.5, *б*. Нетрудно написать в этом случае уравнения (5.10) для граней 1-й зоны Бриллюэна. Полагая, например, $\boldsymbol{b}_g = \boldsymbol{b}_1$, получим из выражений (5.10) и (П.13.3)

$$\frac{2\pi^2}{a^2}(\boldsymbol{i}_0 - \boldsymbol{j}_0 + \boldsymbol{k}_0)^2 + \boldsymbol{k}\,\frac{2\pi}{a}(\boldsymbol{i}_0 - \boldsymbol{j}_0 + \boldsymbol{k}_0) = 0$$

или

$$-k_x + k_y - k_z = 3\pi/a.$$

Это уравнение определяет одну из восьми шестиугольных граней. Если положить $\boldsymbol{b}_g = \boldsymbol{b}_1 + \boldsymbol{b}_2 = \frac{4\pi}{a}\boldsymbol{i}_0$, как это следует из

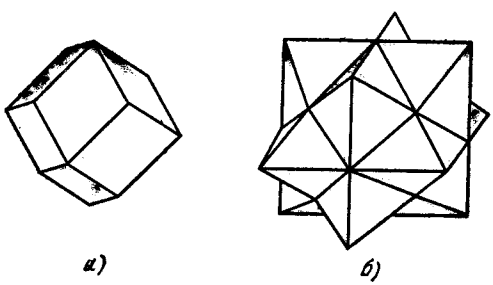


Рис. IV. 6.

(П.13.3), то получим из условия (5.10)

$$k_x = -2\pi/a,$$

что определяет одну из шести квадратных граней.

Аналогично может быть определена форма зон Бриллюэна для прямой решетки объемноцентрированного куба. Так как в этом случае обратная решетка имеет структуру гранецентрированного куба, то первая координационная группа состоит из 12 узлов (гл. I, § 2, п. 2). Двенадцать плоскостей, проведенных в соответствии с уравнением (5.10), полностью определяют в этом случае 1-ю зону Бриллюэна в форме двенадцатигранника (додекаэдра), изображенного на рис. IV.6, *а*. Вторая зона Бриллюэна в этом случае имеет вид, изображенный на рис. IV. 6, *б*.

2. Плоскости в $\boldsymbol{k}$-пространстве, определяемые уравнением (5.10), дают границу зоны (разрыва энергии ε_k электрона)

только при условии $V_g \neq 0$, так как только в этом случае, согласно уравнениям (5.6) и (5.9а), коэффициент a_h и поправка к энергии ε' становятся большими. Если решетка сложная, то условие интерференции электронных волн, рассеиваемых атомами отдельных подрешеток, может привести к исчезновению некоторых V_g. В этом случае энергия ε_k для соответствующей грани зоны Бриллюэна не терпит разрыва, т. е. соответствующая грань фактически не является границей энергетической зоны. Как мы сейчас увидим, ситуация здесь полностью совпадает с той, с которой мы встретились при определении структурного фактора рассеяния рентгеновских лучей (гл. I, § 4, п. 5).

Рассмотрим *сложную* решетку с основными периодами $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$, и пусть положение s атомов ячейки относительно ее начала определяется s векторами $\boldsymbol{r}_n = u_n\boldsymbol{a}_1 + v_n\boldsymbol{a}_2 + w_n\boldsymbol{a}_3$ ($n = 1, 2, \ldots, s$), так что *базис* решетки равен (u_n, v_n, w_n). Если все s атомов одинаковы, то потенциальная энергия электрона аддитивно складывается из энергий взаимодействия с каждой из s подрешеток:

$$V(\boldsymbol{r}) = \sum_n V_1(\boldsymbol{r} - \boldsymbol{r}_n), \tag{6.1}$$

где $|\boldsymbol{r} - \boldsymbol{r}_n|$ — расстояние электрона до n-го атома в нулевой ячейке и $V_1(\boldsymbol{r})$ — периодический потенциал, создаваемый одной подрешеткой. Таким образом,

$$V_1(\boldsymbol{r} - \boldsymbol{r}_n) = \sum_g V_{1g} e^{i\boldsymbol{b}_g(\boldsymbol{r} - \boldsymbol{r}_n)} \tag{6.1а}$$

и, следовательно,

$$V(\boldsymbol{r}) = \sum_g \sum_n V_{1g} e^{-i\boldsymbol{b}_g \boldsymbol{r}_n} e^{i\boldsymbol{b}_g \boldsymbol{r}}. \tag{6.1б}$$

Это может быть записано в следующем виде:

$$V(\boldsymbol{r}) = \sum_g V_g e^{i\boldsymbol{b}_g \boldsymbol{r}}, \tag{6.1в}$$

где

$$V_g = V_{1g} S_g \tag{6.1г}$$

и *структурный множитель*

$$S_g = \sum_{n=1}^{s} e^{-i\boldsymbol{b}_g \boldsymbol{r}_n} = \sum_{n=1}^{s} e^{-2\pi i (g_1 u_n + g_2 v_n + g_3 w_n)}. \tag{6.1д}$$

Выражение для структурного множителя электронов (6.1д) совпадает с определением структурного множителя для рентгеновских лучей (I.4.11а). Таким образом, если $S_g = S_{g_1 g_2 g_3}$ равно нулю, то все соответствующие физически эквивалентные плоскости $\{g_1 g_2 g_3\}$ не эффективны как в рассеянии рентгеновских лучей, так и в возмущении движения почти свободных электронов.

Рассмотрим простую кубическую решетку, для которой обратная решетка тоже простая кубическая, с основными векто-

рами $b_i = 2\pi/a$ $(i = 1, 2, 3)$. В этом случае расстояние плоскости $\{g_1 g_2 g_3\}$ до начала координат, как это видно из уравнения для этой плоскости (5.10), равно

$$h_{g_1 g_2 g_3} = \frac{1}{2} b_{g_1 g_2 g_3} = \frac{\pi}{a} \sqrt{g_1^2 + g_2^2 + g_3^2}.$$

В табл. IV.1 даны для ряда плоскостей $\{g_1 g_2 g_3\}$ их число и их расстояние до начала координат.

Таблица IV.1

Плоскость	{100}	{110}	{111}	{200}	{220}
Число плоскостей	6	12	8	6	12
$h_{g_1 g_2 g_3} \times \frac{a}{\pi}$	1	$\sqrt{2}$	$\sqrt{3}$	2	$2\sqrt{2}$

В качестве первого примера сложной решетки рассмотрим куб с центрированными гранями, который уже был рассмотрен как простая решетка в предыдущем разделе. Если выбрать в качестве a_i ребра куба, то элементарная ячейка гранецентрированного куба содержит четыре атома с базисом $(u_n v_n w_n) =$ $= (000,\ 0\,{}^1/_2\ {}^1/_2,\ {}^1/_2\ 0\,{}^1/_2,\ {}^1/_2\ {}^1/_2\ 0)$. Структурный фактор, согласно условию (6.1д), равен

$$S_{g_1 g_2 g_3} = 1 + e^{-\pi i (g_2 + g_3)} + e^{-\pi i (g_1 + g_3)} + e^{-\pi i (g_1 + g_2)}. \qquad (6.2)$$

Отсюда прямо следует, что

$$S_{100} = 0, \quad S_{110} = 0, \quad S_{111} = 4, \quad S_{200} = 4. \qquad (6.2а)$$

Таким образом, 1-я зона Бриллюэна ограничена восемью плоскостями $\{111\}$ и шестью плоскостями $\{200\}$. При пересечении этих плоскостей образуется усеченный октаэдр или четырнадцатигранник, изображенный на рис. IV.5, *а*. Этот результат совпадает с тем, что было получено в предыдущем пункте.

Рассмотрим теперь действительно сложную решетку типа алмаза (см. рис. I.13), не сводимую к простой решетке Браве (гл. I, § 2, п. 5). Так как решетка типа алмаза может быть представлена как две решетки гранецентрированного куба, сдвинутые друг относительно друга вдоль объемной диагонали куба на ${}^1/_4$ его длины, то, если выбрать элементарную ячейку в виде куба (рис. 1.13), она будет содержать восемь атомов с базисом $(000,\ 0\,{}^1/_2\ {}^1/_2,\ {}^1/_2\ 0\ {}^1/_2,\ {}^1/_2\ {}^1/_2\ 0,\ {}^1/_4\ {}^1/_4\ {}^1/_4,\ {}^1/_4\ {}^3/_4\ {}^3/_4,\ {}^3/_4\ {}^1/_4\ {}^3/_4,\ {}^3/_4\ {}^3/_4\ {}^1/_4)$. Структурный фактор

$$S_{g_1 g_2 g_3} = 1 + e^{-\pi i (g_2 + g_3)} + e^{-\pi i (g_2 + g_3)} + e^{-\pi i (g_1 + g_2)} + e^{-{}^1/_2 \pi i (g_1 + g_2 + g_3)} +$$
$$+ e^{-{}^1/_2 \pi i (g_1 + 3g_2 + 3g_3)} + e^{-{}^1/_2 \pi i (3g_1 + g_2 + 3g_3)} + e^{-{}^1/_2 \pi i (3g_1 + 3g_2 + g_3)}. \qquad (6.3)$$

Отсюда следует, что

$$S_{100}=0,\ S_{110}=0,\ S_{111}=6,\ S_{200}=0,\ S_{211}=0,\ S_{220}=8,\ S_{221}=0.$$

Так как кристаллические решетки германия и кремния можно рассматривать как два сдвинутых друг относительно друга гранецентрированных куба, то 1-я зона Бриллюэна для них имеет форму четырнадцатигранника, изображенного на рис. IV. 5, *а*[1]).

§ 7. Приближение сильно связанных электронов

1. В предыдущем параграфе мы рассматривали периодический потенциал $V(r)$, действующий на электрон как малое возмущение его свободного движения. Только при выполнении определенных интерференционных условий (5.10) движение электрона испытывает сильное возмущение. Такое рассмотрение количественно оправдано только в том случае, когда кинетическая энергия электрона велика по сравнению с пространственными колебаниями его потенциальной энергии $V(\boldsymbol{r})$. Этот случай реализуется, например, при облучении кристалла электронами хотя бы в несколько сот *эв*. С другой стороны, из квантовомеханической теоремы вириала следует, что средняя кинетическая энергия электрона в атоме, молекуле или кристалле должна быть порядка колебаний его потенциальной энергии, поэтому к электронам кристалла неприменимо приближение слабой связи. Можно попытаться подойти к вопросу с другой стороны, считая, что состояние электрона в изолированном атоме мало изменится при образовании из атомов кристалла. Очевидно, что это приближение сильно связанных электронов лучше оправдывается для электронов глубоких энергетических уровней атомов, т. е. для таких, которые относительно слабо взаимодействуют с атомами других узлов решетки. Конечно, приближение ни сильно, ни слабо связанных электронов не описывает правильно с количественной точки зрения состояние электронов в зоне проводимости кристалла. Поэтому оба эти приближения не могут быть использованы для количественных расчетов энергетического спектра и волновых функций электронов проводимости в конкретных кристаллах. Существенно, однако, то, что они дают хорошую иллюстрацию к общим выводам о движении электрона в периодическом поле. В некоторых случаях эти иллюстрации позволяют дать ряд новых качественных выводов о состоянии электрона в периодическом поле.

---

[1]) Другие примеры вычисления структурных факторов и зон Бриллюэна для сложных решеток имеются в книгах: Mott N. F., Jones H. The Theory of Properties of Metals and Alloys.— Oxford, 1936 и Wilson A. H. The Theory of Metals.— Gambridge, 1953.

2. Пусть волновая функция электрона $\psi_0(r)$ в изолированном атоме в s-состоянии удовлетворяет уравнению Шредингера

$$-\frac{\hbar^2}{2m}\nabla^2\psi_0 + \mathcal{U}(r)\psi_0 = \varepsilon_0\psi_0, \tag{7.1}$$

где $\mathcal{U}(r)$ — сферически симметричное поле изолированного иона (металл) или нейтрального атома (полупроводник, когда электрон проводимости «лишний» в кристалле), ε_0 — энергия рассматриваемого состояния электрона в ионе (атоме). Пусть

$$\int \psi_0^*\psi_0\, d\tau = 1, \tag{7.1а}$$

т. е. волновая функция ψ_0 нормирована на единицу.

Вблизи n-го узла решетки электрон в изолированном атоме описывается волновой функцией $\psi_0(|\boldsymbol{r}-\boldsymbol{a}_n|)$. В идеальной решетке все $N = G^3$ узлов основной области полностью эквивалентны, поэтому состояние электрона с энергией ε_0 является N-кратно (без учета спина) вырожденным. Под влиянием взаимодействия электрона со всеми атомами уровень энергии ε_0 расщепится в полосу (зону) и вырождение, во всяком случае частично, будет снято.

Как известно из теории возмущений вырожденных состояний[1]) и, как представляется наглядным из общих соображений, волновая функция нулевого приближения должна быть сконструирована из всех вырожденных волновых функций в виде линейного выражения

$$\psi(\boldsymbol{r}) = \sum_n C_n \psi_0(|\boldsymbol{r}-\boldsymbol{a}_n|). \tag{7.2}$$

Постоянные коэффициенты C_n определяются по известному правилу[2]). Мы воспользуемся для определения коэффициентов C_n простым приемом, потребовав, чтобы (7.2) удовлетворяло общим требованиям, предъявляемым к виду блоховской волновой функции (3.1). Легко видеть, что для этого достаточно положить $C_n = \exp(i\boldsymbol{k}\boldsymbol{a}_n)$, т. е.

$$\psi(\boldsymbol{r}) = \sum_n e^{i\boldsymbol{k}\boldsymbol{a}_n}\psi_0(|\boldsymbol{r}-\boldsymbol{a}_n|). \tag{7.2а}$$

В самом деле, запишем выражение (7.2а) в виде

$$\psi(\boldsymbol{r}) = e^{i\boldsymbol{k}\boldsymbol{r}}\sum_n e^{i(\boldsymbol{k},\ \boldsymbol{a}_n - \boldsymbol{r})}\psi_0(|\boldsymbol{r}-\boldsymbol{a}_n|). \tag{7.2б}$$

Докажем, что множитель, стоящий при $e^{i\boldsymbol{k}\boldsymbol{r}}$, обладает периодичностью решетки, т. е. может рассматриваться как модулирующий множитель $u_k(\boldsymbol{r})$ в блоховской волновой функции (3.1).

---

[1]) Блохинцев Д. И., § 68.
[2]) См. там же.

Подставим в этот множитель вместо $\boldsymbol{r}$ вектор $\boldsymbol{r}+\boldsymbol{a}_m$; получим

$$\sum_n e^{i(\boldsymbol{k},\, \boldsymbol{a}_n-\boldsymbol{r}-\boldsymbol{a}_m)}\psi_0(|\boldsymbol{r}+\boldsymbol{a}_m-\boldsymbol{a}_n|).$$

Перейдем от суммирования по $\boldsymbol{n}$ к суммированию по $\boldsymbol{l}$, положив $\boldsymbol{a}_n-\boldsymbol{a}_m=\boldsymbol{a}_l$, тогда получим

$$\sum_l e^{i(\boldsymbol{k},\, \boldsymbol{a}_l-\boldsymbol{r})}\psi_0(|\boldsymbol{r}-\boldsymbol{a}_l|),$$

что, очевидно, совпадает с исходным выражением (7.2б). Тем самым мы во всяком случае показали, что выбор волновой функции электрона в виде (7.2а) удовлетворяет условиям трансляционной симметрии. Кроме того, очевидно, что ввиду экспоненциально быстрого спадания атомной волновой функции ψ_0 волновая функция электрона в кристалле (7.2а) вблизи n-го узла ведет себя приближенно, как

$$\psi(\boldsymbol{r}) \approx \text{const}\,\psi_0(|\boldsymbol{r}-\boldsymbol{a}_n|),$$

т. е. как атомная функция $\boldsymbol{n}$-го узла.

Если $V(\boldsymbol{r})$ — самосогласованный периодический потенциал, действующий на электрон, причем, конечно, $V(\boldsymbol{r}) \neq \sum_n \mathcal{U}(\boldsymbol{r}-\boldsymbol{a}_n)$, то точная одноэлектронная волновая функция удовлетворяет уравнению

$$\mathcal{H}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V(\boldsymbol{r})\,\psi = \varepsilon\psi, \tag{7.3}$$

где ε — собственное значение энергии электрона, движущегося в кристалле. Умножая обе части равенства (7.3) слева на ψ^* и интегрируя по основной области кристалла, получим

$$\varepsilon = \frac{\int \psi^*\hat{\mathcal{H}}\psi\, d\tau}{\int \psi^*\psi\, d\tau} = \frac{\int \psi^*\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})\right]\psi\, d\tau}{\int \psi^*\psi\, d\tau}. \tag{7.4}$$

Вычислим энергию электрона ε в предположении, что волновая функция $\psi(\boldsymbol{r})$ задается выражением (7.2а). Обозначая $\boldsymbol{r}-\boldsymbol{a}_n=\boldsymbol{\rho}_n$, получим

$$\hat{\mathcal{H}}\psi = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})\right]\sum_n e^{i\boldsymbol{k}\boldsymbol{a}_n}\psi_0(\boldsymbol{\rho}_n) =$$

$$= \varepsilon_0\sum_n e^{i\boldsymbol{k}\boldsymbol{a}_n}\psi_0(\boldsymbol{\rho}_n) + \sum_n e^{i\boldsymbol{k}\boldsymbol{a}_n}[V(\boldsymbol{r})-\mathcal{U}(\boldsymbol{\rho}_n)]\,\psi_0(\boldsymbol{\rho}_n),$$

где было использовано уравнение (7.1) для замены $-\frac{\hbar^2}{2m}\nabla^2\psi_0(\boldsymbol{\rho}_n)$. Подставляя полученное выражение в (7.4) и заменяя ψ^* согласно

(7.2а), получим

$$\varepsilon = \varepsilon_0 + \frac{\sum\limits_m \sum\limits_n e^{i(k,\, a_n - a_m)} \int \psi_0^* (\rho_m) [V(r) - \mathcal{U}(\rho_n)] \psi_0 (\rho_n)\, d\tau}{\sum\limits_m \sum\limits_n e^{i(k,\, a_n - a_m)} \int \psi_0^* (\rho_m) \psi_0 (\rho_n)\, d\tau}.$$

Ввиду эквивалентности всех узлов числитель и знаменатель не зависят от m и n в отдельности, а только от их разности, т. е. от относительного положения узлов. Поэтому можно положить $m = 0$ (т. е. $a_m = 0$, $\rho_m = r$) и суммирование по m как в числителе, так и в знаменателе заменить умножением на число узлов основной области N. Таким образом,

$$\varepsilon = \varepsilon_0 + \frac{\sum\limits_n e^{ika_n} \int \psi_0^* (r) [V(r) - \mathcal{U}(\rho_n)] \psi_0 (\rho_n)\, d\tau}{\sum\limits_n e^{ika_n} \int \psi_0^* (r) \psi_0 (\rho_n)\, d\tau}. \qquad (7.5)$$

Будем считать, что атомные волновые функции ψ_0 спадают так быстро, что можно пренебречь их перекрытием даже для соседних узлов, т. е.

$$\int \psi_0^* (r) \psi_0 (\rho_n)\, d\tau = \delta_{0n} = \begin{cases} 0, & n \neq 0, \\ 1, & n = 0. \end{cases} \qquad (7.6)$$

Обозначим интеграл, стоящий в числителе дроби (7.5), для $n = 0$

$$\int \psi_0^* (r) [V(r) - \mathcal{U}(r)] \psi_0 (r)\, d\tau = \int |\psi_0 (r)|^2 [V(r) - \mathcal{U}(r)]\, d\tau = -C < 0. \qquad (7.7)$$

Отрицательное значение интеграла (7.7) может быть до некоторой степени обосновано следующим образом. На рис. IV.7 сплошной линией представлен потенциал изолированного атома $\mathcal{U}$, а пунктирной — самосогласованный периодический потенциал V. Если сделать довольно естественное предположение, что взаимодействие между атомами снижает потенциальные барьеры для электрона, как это показано на рисунке, то квадратная скобка в интеграле (7.7) повсюду отрицательна и, следовательно, сам интеграл отрицателен.

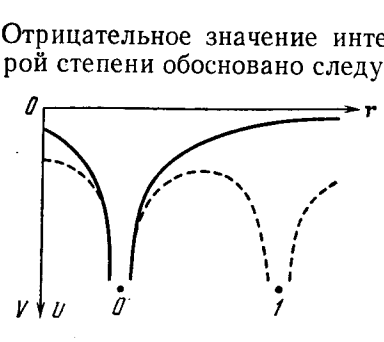


Рис. IV. 7.

Хотя мы пренебрегаем перекрытием самих атомных волновых функций разных узлов, будет последовательным учесть интегралы в числителе дроби (7.5) при $n \neq 0$ для соседних узлов n_0:

$$\int \psi_0^* (r) [V(r) - \mathcal{U}(\rho_{n_0})] \psi_0 (\rho_{n_0})\, d\tau = -A_{n_0}. \qquad (7.8)$$

В самом деле, хотя $\psi_0(\boldsymbol{\rho}_{n_0})$ мало́ вблизи нулевого узла, но это частично компенсируется в этой области большим значением разности $V(\boldsymbol{r})-\mathcal{U}(\boldsymbol{\rho}_{n_0})$. Некоторые общие соображения, на которых мы останавливаться не будем, позволяют предполагать, что и в случае (7.8) интеграл отрицателен. Из соотношения (7.5)—(7.8) получим

$$\varepsilon=\varepsilon_0-C-\sum_{n_0}A_{n_0}e^{ika_{n_0}}. \tag{7.9}$$

Если волновая функция $\psi_0(\boldsymbol{r})$ относится к s-состоянию, то A_{n_0} зависит только от расстояния между нулевым узлом и атомом $\boldsymbol{n}_0$ первой координатной группы, т. е. одинаково для всех этих атомов. В этом случае

$$\varepsilon=\varepsilon_0-C-A\sum_{n_0}e^{ika_{n_0}}. \tag{7.10}$$

В простой кубической решетке каждый атом окружен шестью ближайшими соседями. Направив оси x, y и z по ребрам куба, получим

$$\sum_{n_0}^{1,6}e^{ika_{n_0}}=e^{ikxa}+e^{-ikxa}+e^{ikya}+e^{-ikya}+e^{ikza}+e^{-ikza}=$$
$$=2\left[\cos ak_x+\cos ak_y+\cos ak_z\right],$$

где a — постоянная решетки. Подставляя эти результаты в (7.10), получим

$$\varepsilon=\varepsilon_0-C-2A\left[\cos ak_x+\cos ak_y+\cos ak_z\right]. \tag{7.10а}$$

В случае объемноцентрированного куба каждый атом окружен восемью ближайшими соседями, и вместо уравнения (7.10а) мы получим (Приложение 14)

$$\varepsilon=\varepsilon_0-C-8A\cos\frac{ak_x}{2}\cos\frac{ak_y}{2}\cos\frac{ak_z}{2}. \tag{7.10б}$$

Аналогично может быть вычислена энергия электрона в решетке гранецентрированного куба:

$$\varepsilon=\varepsilon_0-C-4A\left[\cos\frac{ak_x}{2}\cos\frac{ak_y}{2}+\cos\frac{ak_x}{2}\cos\frac{ak_z}{2}+\cos\frac{ak_y}{2}\cos\frac{ak_z}{2}\right]. \tag{7.10в}$$

Из выражений (7.10а)—(7.10в) видно, что при образовании кристалла из отдельных атомов энергия электрона в изолированном атоме ε_0 в результате взаимодействия с соседними атомами смещается на величину C и расщепляется в энергетическую зону, в пределах которой энергия электрона периодически зависит от составляющих волнового вектора k. Такому расщеплению подвергается каждый стационарный энергетический уровень изолированного атома; поэтому качественно картина будет подобна той, которая изображена на рис. IV.2 для почти

свободных электронов. Ширина разрешенной энергетической зоны, как мы увидим ниже, пропорциональна величине A, т. е. в большой мере определяется перекрытием атомных волновых функций $\psi_0(r)$ соседних атомов. Для внешних валентных электронов атомов, которые нас обычно и интересуют, это перекрытие велико, так что ширина энергетической зоны достигает нескольких *эв*, т. е. порядка и даже превосходит расстояние между уровнями энергии изолированного атома. Строго говоря, это означает, что развитый выше метод сильной связи неприменим к этому случаю. Для электронов внутренних оболочек атомов это расщепление мало́: так, для K-электронов в решетке металлического натрия оно порядка $2\cdot10^{-19}$ *эв*, так что уровень практически остается резким. В результате спектр энергии электрона в кристалле приобретает вид, изображенный на рис. IV.8, где разрешенные полосы энергии заштрихованы.

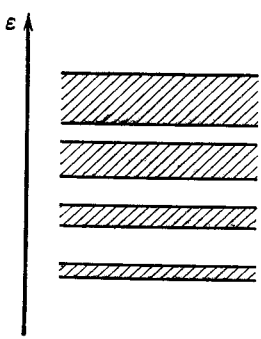


Рис. IV. 8.

Как можно представить себе движение электрона сквозь решетку с точки зрения приближения сильной связи? В изолированном атоме электрон пребывает на стационарном уровне ε_0 неопределенно долгое время. При сближении таких одинаковых атомов и образовании решетки электрон приобретает возможность посредством квантовомеханического туннельного эффекта перейти от одного узла решетки к соседнему. При этом резкий уровень энергии ε_0 из-за взаимодействия атомов расширяется в полосу шириной ε_{ab}, которая связана с временем пребывания электрона вблизи данного узла τ соотношением неопределенности $\varepsilon_{ab}\tau \approx \hbar$. Для внешних электронов $\varepsilon_{ab} \approx 10$ *эв* и соответствующие $\tau \approx 10^{-15}$ *сек*, но для K-электронов натрия, когда $\varepsilon_{ab} \approx 2\cdot10^{-19}$ *эв*, электрон переходит от одного узла к соседнему в среднем за время $\tau \approx 1$ час. Однако даже в этом последнем случае в стационарном состоянии электрон распределен с одинаковой вероятностью по всем узлам решетки кристалла.

3. Проиллюстрируем свойства электрона, движущегося в идеальном периодическом поле (§ 3), на примере сильно связанных электронов в простой кубической решетке. Будем энергию электрона ε в выражениях (7.10) отсчитывать от уровня $\varepsilon_0 - C$, т. е. положим $\varepsilon_0 - C = 0$. Из (7.10а) видно, что энергия электрона ε имеет при $k_x = k_y = k_z = 0$, т. е. в центре бриллюэновской зоны, минимальное значение, равное $\varepsilon_b = -6A$, а при $k_x = k_y = k_z = \pi/a$, т. е. в вершинах куба зоны, — максимальное значение $\varepsilon_a = +6A$. Таким образом, в случае простой кубической решетки ширина зоны $\varepsilon_{ab} = \varepsilon_a - \varepsilon_b = 12A$.

Для малых $\boldsymbol{k}$, т. е. $ak_x \ll 1$, $ak_y \ll 1$ и $ak_z \ll 1$, получим, разлагая косинусы в ряд,

$$\varepsilon = -2A\left[\left(1-\frac{(ak_x)^2}{2}\right)+\left(1-\frac{(ak_y)^2}{2}\right)+\left(1-\frac{(ak_z)^2}{2}\right)\right] = \\ = -6A + Aa^2(k_x^2+k_y^2+k_z^2) = \varepsilon_b + Aa^2k^2. \quad (7.11)$$

Таким образом, для малых $\boldsymbol{k}$ энергия электрона, так же как и в случае слабой связи, не зависит от направления $\boldsymbol{k}$ и пропорциональна k^2. Скалярная эффективная масса электрона у нижнего края зоны

$$m_n^*(b) = \frac{\hbar^2}{\left(\frac{\partial^2\varepsilon}{\partial k_x^2}\right)_b} = \frac{\hbar^2}{2Aa^2} = \frac{6\hbar^2}{a^2\varepsilon_{ab}}, \quad (7.12)$$

т. е. положительна, и, если приближенно считать $a = \text{const}$ для разных кристаллов, обратно пропорциональна ширине зоны ε_{ab}. У верхнего края зоны положим $\frac{\pi}{a} - k_x = k_x'$ и т. д. и будем считать $ak_x' \ll 1$ и т. д., тогда

$$\varepsilon = -2A[\cos(\pi - ak_x') + \cos(\pi - ak_y') + \cos(\pi - ak_z')] = \\ = 2A[\cos ak_x' + \cos ak_y' + \cos ak_z'] = \\ = 6A - Aa^2(k_x'^2 + k_y'^2 + k_z'^2) = \varepsilon_a - Aa^2k'^2. \quad (7.11а)$$

Аналогично эффективная масса электрона у верхнего края зоны

$$m_n^*(a) = \frac{\hbar^2}{\left(\frac{\partial^2\varepsilon}{\partial k_x'^2}\right)_a} = -\frac{\hbar^2}{2Aa^2}, \quad (7.12а)$$

т. е. отрицательна. Следовательно, эффективная масса дырки

$$m_p^*(a) = -m_n^*(a) = \hbar^2/2Aa^2, \quad (7.12б)$$

т. е. совпадает с эффективной массой электрона нижнего края зоны (7.12).

Очевидно, что изоэнергетические поверхности в $\boldsymbol{k}$-пространстве, вблизи центра зоны и у вершин куба ($\pm k_x = \pm k_y = \pm k_z = \pi/a$) имеют сферическую форму. При промежуточных значениях энергии они имеют более сложную форму, показанную на рис. IV.9, где (*а*) соответствует энергия $\varepsilon(\boldsymbol{k}) = -2A$, а (*б*) — энергия $\varepsilon(\boldsymbol{k}) = 0$ (мы по-прежнему полагаем $\varepsilon_0 - C = 0$).

В силу периодичности энергии ε в $\boldsymbol{k}$-пространстве с периодами $\boldsymbol{b}_i$ форма поверхности $\varepsilon = \text{const}$ повторяется во всех элементарных ячейках (увеличенных в масштабе 2π) пространства обратной решетки. В случае, изображенном на рис. IV.9, *б*, многосвязная поверхность постоянной энергии имеет вид, изображенный на рис. IV.10. В случае поверхности постоянной энергии

$\varepsilon(\boldsymbol{k}) = -2A$ (рис. IV.9, *а*) поверхности энергии соседних ячеек пространства обратной решетки имеют только общие точки соприкосновения в центре обратных граней бриллюэновской зоны.

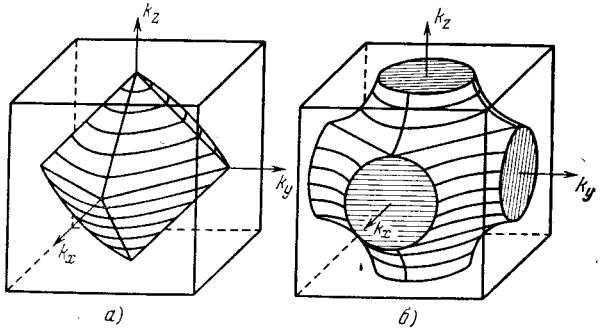


Рис. IV.9.

При энергии электрона, меньшей $\varepsilon = -2A$, замкнутые поверхности постоянной энергии в разных ячейках обратной решетки не будут иметь общих точек соприкосновения.

Эта топология поверхностей постоянной энергии для $\varepsilon(\boldsymbol{k})$, равной энергии ферми-электронов проводимости металла, играет

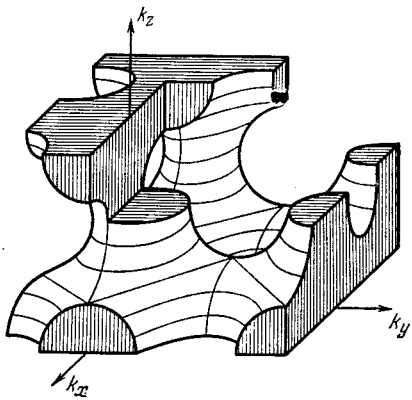


Рис. IV.10.

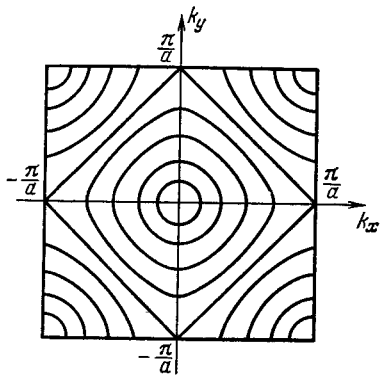


Рис. IV.11.

важную роль при исследовании гальваномагнитных явлений в металлах в сильном магнитном поле (И. М. Лифшиц).

На рис. IV.11 представлены изоэнергетические линии на плоскости $k_z = 0$. Изоэнергетические отрезки типа $k_x + k_y = \pi/a$ соответствуют энергии

$$\varepsilon = -2A\left[\cos ak_x + \cos(\pi - ak_x) + 1\right] = -2A.$$

Мы видим, что при энергиях, далеких от краев зоны, поведение сильно связанных электронов значительно отличается от почти свободных.

Из выражений (3.32) и (7.10а) следует, что средняя квантовомеханическая скорость электронов в $\boldsymbol{k}$-состоянии равна

$$\boldsymbol{v}=\frac{2Aa}{\hbar}[\sin ak_x\boldsymbol{i}_0+\sin ak_y\boldsymbol{j}_0+\sin ak_z\boldsymbol{k}_0], \tag{7.13}$$

т. е. зависит не только от абсолютной величины волнового вектора $\boldsymbol{k}$, но и от его направления. Из (7.13) видно, что у нижнего и верхнего краев зоны скорость равна нулю. Для малых $\boldsymbol{k}$

$$\boldsymbol{v}=\frac{2Aa^2}{\hbar}(k_x\boldsymbol{i}_0+k_y\boldsymbol{j}_0+k_z\boldsymbol{k}_0)=\frac{2Aa^2}{\hbar}\boldsymbol{k}=\frac{\boldsymbol{p}}{m_n^*(b)}, \tag{7.13а}$$

где квазиимпульс $\boldsymbol{p}=\hbar\boldsymbol{k}$, а эффективная масса $m_n^*(b)$ равна (7.12).

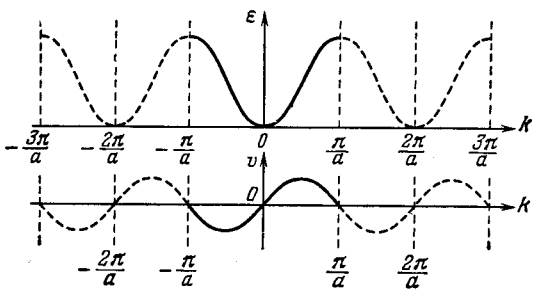


Рис. IV. 12.

Полагая $k_y=k_z=0$ и $k_x=k$, получим для движения вдоль x

$$\varepsilon=-2A\cos ak, \tag{7.14}$$

$$v=\frac{2Aa}{\hbar}\sin ak. \tag{7.14а}$$

На рис. IV.12 представлена зависимость ε и v от k согласно (7.14) и (7.14а). В согласии с (4.4)

$$\int\limits_{-\pi/a}^{+\pi/a}v(k)\,dk=\frac{2Aa}{\hbar}\int\limits_{-\pi/a}^{+\pi/a}\sin ak\,dk=0,$$

т. е. средняя скорость электрона (ток) по всей зоне равна нулю. На рис. IV.12 особенно ясно видна периодичность ε и v в зависимости от волнового вектора k и возможность рассмотрения их в пределах приведенной зоны:

$$-\frac{\pi}{a}\leqslant k\leqslant+\frac{\pi}{a}.$$

Плотность состояний на 1 $см^3$, согласно выражениям (3.26) и (7.10а), равна

$$g(\varepsilon)=\frac{1}{8\pi^3}\int\frac{d\sigma}{|\operatorname{grad}_{\boldsymbol{k}}\varepsilon|}=\frac{1}{16\pi^3 Aa}\int\frac{d\sigma}{\sqrt{\sin^2 ak_x+\sin^2 ak_y+\sin^2 ak_z}}, \qquad (7.15)$$

где интегрирование ведется по поверхности

$$\varepsilon=-2A(\cos ak_x+\cos ak_y+\cos ak_z)=\text{const}. \qquad (7.15а)$$

Используя это соотношение, можно привести интеграл (7.15) к одномерному и вычислить его приближенно. Кривая *а* на рис. IV.13 представляет результат такого вычисления. Заметим, что кривая симметрична относительно точки $\varepsilon_{ab}/2$. Кривая *б* на том же рисунке изображает $g(\varepsilon)\sim\sqrt{\varepsilon}$ для свободных электронов, но учитывает, начиная с некоторого ε, уменьшение числа состояний из-за пересечения сферической поверхности энергии с гранями кубической бриллюэновской зоны.

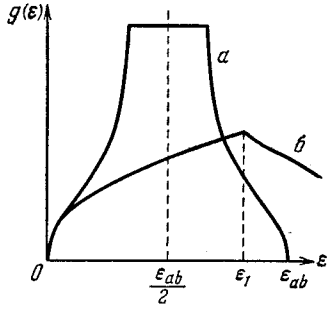


Рис. IV. 13.

4. Для объемноцентрированной кубической решетки 1-я зона Бриллюэна имеет форму правильного додекаэдра, изображенного на рис. IV.6, *а*. Из соотношения (7.10б) видно, что энергия электрона ε имеет минимальное значение[1]) $\varepsilon_b=-8A$ в центре зоны при $\boldsymbol{k}=0$ и максимальное $\varepsilon_a=+8A$ в точке $k_x=2\pi/a$, $k_y=k_z=0$ и в пяти других эквивалентных точках в центрах квадратных граней зоны. Таким образом, ширина зоны $\varepsilon_{ab}=\varepsilon_a-\varepsilon_b=16A$, т. е. и в этом случае пропорциональна интегралу перекрытия A. Плоскости $k_x=\pi/a$ и пяти другим эквивалентным плоскостям соответствует постоянная энергия $\varepsilon=0$. На рис. IV.14 представлены изоэнергетические линии, образованные сечением бриллюэновской зоны плоскостями: *а*) $k_z=0$ и *б*) $k_z=\pi/2a$.

Аналогично тому, как это было сделано в предыдущем разделе для простой кубической решетки, можно получить разложение энергии ε у нижнего и верхнего краев энергетической зоны, определить эффективные массы, среднюю скорость электрона $\boldsymbol{v}$ и плотность состояний $g(\varepsilon)$.

Для гранецентрированной кубической решетки 1-я зона Бриллюэна имеет форму усеченного октаэдра, изображенного на рис. IV.5, *а*. Из уравнения (7.10в) следует, что минимальная энергия при $\boldsymbol{k}=0$ равна $\varepsilon_b=-12A$, а максимальная энергия $\varepsilon_a=4A$ для $k_x=2\pi/a$, $k_y=k_z=0$ и пяти других эквивалентных

[1]) Мы опять полагаем $\varepsilon_0-C=0$.

точек, так что ширина зоны опять $\varepsilon_{ab} = 16A$. В центрах каждой из восьми шестиугольных граней ($\pm k_x = \pm k_y = \pm k_z = \pi/a$) энергия $\varepsilon = 0$. На рис. IV.15 представлены для этого случая изоэнергетические поверхности, причем *а*) соответствует меньшему, а *б*)—

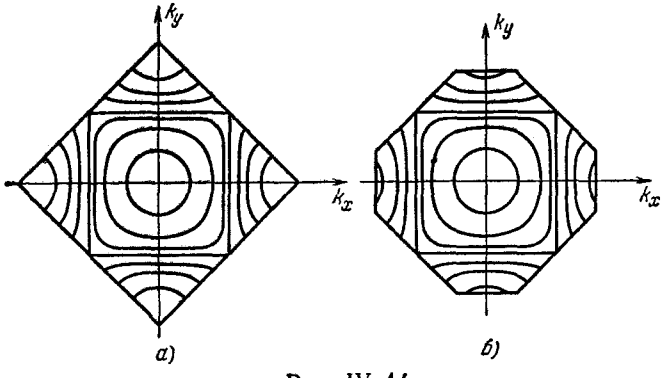


Рис. IV.14.

большему значению энергии ε. На рис. IV.16 даны изоэнергетические линии, полученные при сечении бриллюэновской зоны плоскостью $k_z = 0$.

Конечно, и в случае гранецентрированной решетки можно рассмотреть разложение энергии ε вблизи краев энергетической

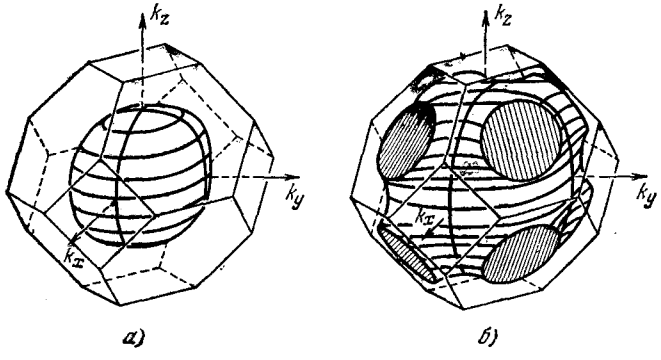


Рис. IV.15.

зоны, эффективные массы и т. д. Результаты, которые при этом получаются, не отличаются качественно от тех, которые имеют место для простой и объемноцентрированной кубических решеток.

5. Рассмотрим в приближении сильной связи поведение электрона в простой кубической решетке в том случае, когда он в изолированном атоме (ионе) находится в *p*-состоянии.

Волновая функция электрона в произвольном сферически симметричном поле $\mathcal{U}(r)$ имеет вид [1])

$$\psi_{nlm}(r, \vartheta, \varphi) = R_{nl}(r)\, Y_{lm}(\vartheta, \varphi),$$

где $R_{nl}(r)$ — радиальная часть волновой функции, зависящая от главного квантового числа n и орбитального квантового числа $l \leqslant n-1$, а $Y_{lm}(\vartheta, \varphi)$ — шаровая функция, описывающая угловую зависимость волновой функции, зависящая от l и магнитного квантового числа $m = 0, \pm 1, \pm 2, \ldots, \pm l$. Энергия электрона зависит только от квантовых чисел n и l (в кулоновском поле — только от n); поэтому p-состоянию электрона с $l = 1$ соответствуют три вырожденные волновые функции с $m = 0, +1, -1$:

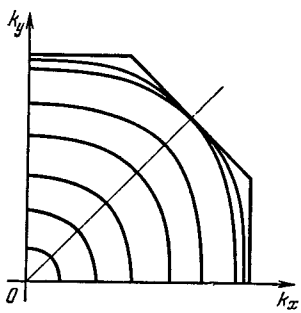


Рис. IV. 16.

$$\psi_{n10} = R_{n1}(r)\left(\frac{3}{4\pi}\right)^{1/2}\cos\vartheta,$$

$$\psi_{n11} = R_{n1}(r)\left(\frac{3}{8\pi}\right)^{1/2}\sin\vartheta e^{i\varphi}, \quad (7.16)$$

$$\psi_{n1,-1} = R_{n1}(r)\left(\frac{3}{8\pi}\right)^{1/2}\sin\vartheta e^{-i\varphi}.$$

В силу линейности и однородности уравнения Шредингера всякая линейная комбинация волновых функций (7.16) тоже является волновой функцией p-состояния электрона. В дальнейшем удобно использовать в качестве волновых функций три следующие линейные комбинации (7.16):

$$\frac{1}{2}[\psi_{n11} + \psi_{n1,-1}] = R_{n1}(r)\left(\frac{3}{8\pi}\right)^{1/2}\sin\vartheta\cos\varphi = xf(r) \equiv \psi_x(\boldsymbol{r}),$$

$$\frac{1}{2i}[\psi_{n11} - \psi_{n1,-1}] = R_{n1}(r)\left(\frac{3}{8\pi}\right)^{1/2}\sin\theta\sin\varphi = yf(r) \equiv \psi_y(\boldsymbol{r}), \quad (7.16a)$$

$$\frac{2}{\sqrt{2}}\psi_{n10} = R_{n1}(r)\left(\frac{3}{8\pi}\right)^{1/2}\cos\vartheta = zf(r) \equiv \psi_z(\boldsymbol{r}).$$

Здесь $x = r\sin\vartheta\cos\varphi$, $y = r\sin\vartheta\sin\varphi$ и $z = r\cos\vartheta$ — прямоугольные координаты электрона, а $f(r)$ — некоторая радиально-симметричная функция, зависящая от вида потенциала $\mathcal{U}(r)$.

Волновая функция электрона в кристалле может быть записана аналогично (7.2а). Однако мы должны теперь учесть не только трансляционное вырождение электрона, учитываемое в (7.2а) множителем $\exp i(\boldsymbol{k}\boldsymbol{a}_n)$ и суммированием по всем эквивалентным узлам решетки $\boldsymbol{n}$, но и вырождение (7.16). Таким образом, положим волновую функцию электрона в кристалле

$$\psi(\boldsymbol{r}) = \sum_{\boldsymbol{n}} e^{i\boldsymbol{k}\boldsymbol{a}_n}\{\alpha\psi_x(\boldsymbol{r}-\boldsymbol{a}_n) + \beta\psi_y(\boldsymbol{r}-\boldsymbol{a}_n) + \gamma\psi_x(\boldsymbol{r}-\boldsymbol{a}_n)\}, \quad (7.17)$$

[1]) Блохинцев Д. И., § 49.

где коэффициенты α, β и γ должны быть определены из общей теории возмущения вырожденных состояний.

Можно показать (Приложение 15), что в простом кубическом кристалле состояния ψ_x, ψ_y и ψ_z не комбинируют друг с другом, т. е. только один из коэффициентов α, β и γ может быть отличен от нуля. Таким образом, при заданном значении волнового вектора $\boldsymbol{k}$ электрону в кристалле соответствуют три волновые функции, для каждой из которых имеет место своя зависимость энергии ε от $\boldsymbol{k}$.

Энергетические зоны, соответствующие этим трем волновым функциям, в точности перекрываются, так что в простом кубическом кристалле вырождение p-уровня не снимается.

Расчет показывает (Приложение 15), что для $\alpha \neq 0$, $\beta = \gamma = 0$ энергия электрона

$$\varepsilon_1 = \varepsilon_0 - C + 2A\cos ak_x - 2B(\cos ak_y + \cos ak_z). \qquad (7.18)$$

Для $\beta \neq 0$, $\alpha = \gamma = 0$ энергия электрона

$$\varepsilon_2 = \varepsilon_0 - C + 2A\cos ak_y - 2B(\cos ak_x + \cos ak_z) \qquad (7.18а)$$

и для $\gamma \neq 0$, $\alpha = \beta = 0$ энергия электрона

$$\varepsilon_3 = \varepsilon_0 - C + 2A\cos ak_z - 2B(\cos ak_x + \cos ak_y). \qquad (7.18б)$$

Здесь ε_0 — энергия p-электрона в изолированном атоме,

$$C = -\int \psi_x^2(\boldsymbol{r})\,[V(\boldsymbol{r}) - \mathcal{U}(\boldsymbol{r})]\,d\tau \qquad (7.19)$$

и

$$\int \psi_x(\boldsymbol{r})\,\psi_x(\boldsymbol{r}-\boldsymbol{a}_{n_0})\,[V(\boldsymbol{r}) - \mathcal{U}(|\boldsymbol{r}-\boldsymbol{a}_{n_0}|)]\,d\tau = \begin{cases} A, \\ -B, \end{cases} \qquad (7.19а)$$

где интеграл равен A, когда соседний атом $\boldsymbol{n}_0$ взят по оси x, и равен $-B$, когда соседний атом $\boldsymbol{n}_0$ взят по оси y или z. В силу кубической симметрии интегралы (7.19) и (7.19а) не меняются от замены ψ_x на ψ_y или ψ_z.

Потенциальные энергии $V(\boldsymbol{r})$ и $\mathcal{U}(\boldsymbol{r})$ имеют тот же смысл, что и в выражениях (7.7) и (7.8). Знаки постоянных A и B не могут быть установлены вполне однозначно. Мы рассмотрим два случая: I. $A > 0$ и $B > 0$ и II. $A > 0$ и $B < 0$. В дальнейшем мы опять будем энергию электрона во всех трех состояниях отсчитывать от уровня $\varepsilon_0 - C$, т. е. положим $\varepsilon_0 - C = 0$.

I случай: $A > 0$, $B > 0$. Для всех трех состояний электрона (7.18), (7.18а), (7.18б) минимум энергии равен

$$\varepsilon_{1b} = \varepsilon_{2b} = \varepsilon_{3b} = -2A - 4B. \qquad (7.20а)$$

В каждом из трех состояний два эквивалентных минимума расположены в центрах двух противоположных квадратных граней зоны Бриллюэна. Максимальное значение энергии для всех трех

состояний тоже одинаково и равно

$$\varepsilon_{1a} = \varepsilon_{2a} = \varepsilon_{3a} = +2A + 4B, \tag{7.20а}$$

так что ширина зоны трех перекрывающихся полос энергии равна

$$\varepsilon_{ab} = \varepsilon_{1a} - \varepsilon_{1b} = 4A + 8B. \tag{7.20б}$$

Максимумы для каждого из трех состояний расположены в центрах четырех ребер куба, перпендикулярных к граням, на которых располагаются соответствующие минимумы. Таким образом, в зоне Бриллюэна имеется всегда 6 минимумов и 12 максимумов.

Разложим энергию ε_1 (7.18) в ряд вблизи минимума $k_x = \pi/a$, $k_y = k_z = 0$, по малым величинам $k'_x = \frac{\pi}{a} - k_x$, k_y и k_z:

$$\begin{aligned}\varepsilon_1 &= 2A\cos(\pi - ak'_x) - 2B(\cos ak_y + \cos ak_z) = \\ &= -2A\cos ak'_x - 2B(\cos ak_y + \cos ak_z) = \\ &= -2A\left[1 - \frac{(ak'_x)^2}{2}\right] - 2B\left[\left(1 - \frac{(ak_y)^2}{2}\right) + \left(1 - \frac{(ak_z)^2}{2}\right)\right] = \\ &= \varepsilon_{1b} + Aa^2k'^2_x + Ba^2(k_y^2 + k_z^2). \end{aligned} \tag{7.21}$$

Мы видим, что вблизи этого минимума изоэнергетические поверхности — эллипсоиды вращения вокруг оси x. Компоненты тензора эффективной массы

$$\begin{aligned} m_{xx}^{-1} &= \frac{1}{\hbar^2}\left(\frac{\partial^2\varepsilon_1}{\partial k_x'^2}\right)_{\min} = \frac{2Aa^2}{\hbar^2} \equiv \frac{1}{m_1}, \\ m_{yy}^{-1} &= m_{zz}^{-1} = \frac{1}{\hbar^2}\left(\frac{\partial^2\varepsilon_1}{\partial k_y^2}\right)_{\min} = \frac{2Ba^2}{\hbar^2} \equiv \frac{1}{m_2}. \end{aligned} \tag{7.21а}$$

Из выражений (7.21) и (7.21а)

$$\varepsilon_1 - \varepsilon_{1b} = \frac{\hbar^2k'^2_x}{2m_1} + \frac{\hbar^2k_y^2 + \hbar^2k_z^2}{2m_2}. \tag{7.21б}$$

Такие же разложения имеют место вблизи пяти других точек минимума энергии. Таким образом, хотя каждый эллипсоид в отдельности не обладает кубической симметрией, их совокупность удовлетворяет таковой. Аналогичные разложения энергии электрона возможны вблизи точек максимума. В плоскости $k_z = 0$ энергии электрона равны

$$\varepsilon_1 = 2A\cos ak_x - 2B\cos ak_y - 2B, \tag{7.22}$$

$$\varepsilon_2 = 2A\cos ak_y - 2B\cos ak_x - 2B, \tag{7.22а}$$

$$\varepsilon_3 = -2B(\cos ak_x + \cos ak_y) + 2A. \tag{7.22б}$$

На рис. IV.17, *а*, *б*, и *в* представлены изоэнергетические линии, соответствующие ε_1, ε_2 и ε_3. Буквами m и M отмечены точки

минимума и максимума энергии. Для *а* и *б* центру зоны ($k_x = k_y = 0$) соответствует седлообразная точка поверхности энергии

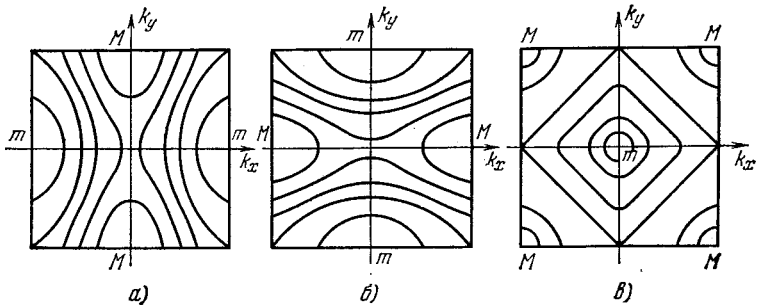


Рис. IV. 17.

$\varepsilon(k_x, k_y)$. На рис. IV.18 представлена зависимость энергии $\varepsilon_1(k_x, k_y)$ вдоль осей k_x и k_y.

II случай: $A > 0$ и $B < 0$. Обозначая B через $-B$ и полагая опять $\varepsilon_0 - C = 0$, получим вместо равенств (7.18)

$$\varepsilon_1 = 2A\cos ak_x + 2B(\cos ak_y + \cos ak_z), \quad (7.23)$$

$$\varepsilon_2 = 2A\cos ak_y + 2B(\cos ak_x + \cos ak_z), \quad (7.23а)$$

$$\varepsilon_3 = 2A\cos ak_z + 2B(\cos ak_x + \cos ak_y). \quad (7.23б)$$

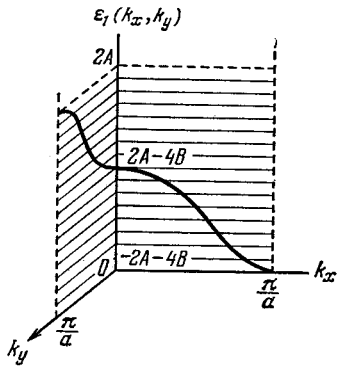


Рис. IV. 18.

Отсюда видно, что минимуму энергии, равному (7.20), соответствуют во всех трех состояниях точки $\pm k_x = \pm k_y = \pm k_z = \pi/a$, т. е. вершины куба зоны Бриллюэна. Максимум энергии, одинаковый во всех трех состояниях, достигается в центре зоны при $k_x = k_y = k_z = 0$. Таким образом, общая ширина энергетической зоны такая же, как и в предыдущем случае (7.20б). Разлагая энергии электрона (7.23а), (7.23б) в ряд по k_x, k_y, k_z вблизи общего максимума, получим

$$\varepsilon_1 = (2A + 4B) - Aa^2k_x^2 - Ba^2(k_y^2 + k_z^2), \quad (7.24)$$

$$\varepsilon_2 = (2A + 4B) - Aa^2k_y^2 - Ba^2(k_x^2 + k_z^2), \quad (7.24а)$$

$$\varepsilon_3 = (2A + 4B) - Aa^2k_z^2 - Ba^2(k_x^2 + k_y^2). \quad (7.24б)$$

Таким образом, изоэнергетические поверхности — одинаковые эллипсоиды с осями вращения, направленными вдоль x, y и z

(рис. IV.19). Если зона почти заполнена электронами, то спектр (7.24), (7.24а), (7.24б), взятый с обратным знаком, соответствует дыркам. Полагая $k_y = k_z = 0$, видим, что

$$\varepsilon_1 = \varepsilon_M - Aa^2k_x^2, \tag{7.25}$$

$$\varepsilon_2 = \varepsilon_3 = \varepsilon_M - Ba^2k_x^2. \tag{7.25а}$$

На рис. IV.20 представлена зависимость энергий ε_1, ε_2, ε_3 от k_x в предположении, что $A < B$ и, следовательно, кривая 2 соответствует дважды вырожденному состоянию ($\varepsilon_2 = \varepsilon_3$).

В направлении k_y, т. е. при $k_x = k_z = 0$, ε_1 и ε_2 подобны (7.25а),

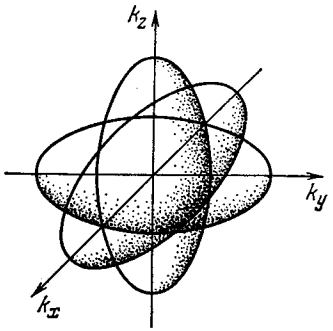


Рис. IV. 19.

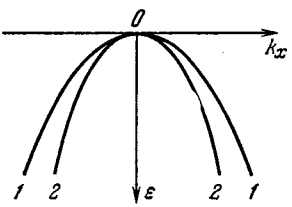


Рис. IV. 20.

а ε_2 подобно (7.25). Состояние в центре бриллюэновской зоны ($\boldsymbol{k} = 0$) является шестикратно вырожденным (с учетом спина).

Если не учитывать спин-орбитального взаимодействия, то такой энергетический спектр наблюдался бы для дырок валентной зоны германия. Учет спин-орбитального взаимодействия приводит к частичному снятию вырождения; отщепившееся двукратно вырожденное состояние смещается в сторону более низких значений энергий, а четырехкратно вырожденное — дает две ветви — легких и тяжелых дырок.

§ 8. Структура энергетических зон и симметрия волновых функций в простой кубической решетке и в кристалле сурмянистого индия

1. Пространственная группа простой кубической решетки — симморфна, поэтому неприводимые представления группы волнового вектора $G_{\boldsymbol{k}}$ для нее могут быть классифицированы по неприводимым представлениям $\Gamma(R)$ — точечной группы, соответствующей волновому вектору $\boldsymbol{k}$ (гл. II, § 9, п. 4).

Так как результаты исследования спектра колебаний в простой кубической решетке в гл. III, § 8, основаны только на соображениях симметрии, то они могут быть без изменения перенесены на свойства электронного спектра в кубическом кристалле. Раз-

ница заключается только в том, что в случае колебаний мы должны центру бриллюэновской зоны Г приписать неприводимое представление Γ_{15} (III.8.17), в то время как в случае электронного спектра выбор состояния электрона в точке Г в широких пределах произволен. Но если мы предположим, что состояние электрона в центре бриллюэновской зоны тоже определяется неприводимым представлением Γ_{15}, то все выводы, полученные для колебательного спектра в гл. III, § 8, остаются в силе и для электронного, надо только заменить понятие «ветвь колебаний» на «зону энергий».

Так, например, трехкратно вырожденное состояние электрона Γ_{15} при переходе к точке Δ (рис. III.13) расщепляется на невырожденное состояние Δ_1 и дважды вырожденное состояние Δ_5 (III.8.19а). Также интерпретируются для электронного спектра выражения (III.8.20), (III.8.21) и (III.8.22). Соотношения совместности, сведенные в табл. III.5, позволяют также просто рассмотреть другие электронные состояния в центре Г и их расщепление при смещении вдоль осей Δ, Λ, Σ [1]).

2. Составим табл. IV.2 преобразования прямоугольных координат x, y, z при действии на них элементов группы тетраэдра $\boldsymbol{T_d}$ (гл. II, § 3). Соответствующие элементы группы $\boldsymbol{T_d}$ обозначены в таблице через $R_i(\ldots)$ $(i=1, 2, 3, \ldots, 24)$, причем в скобках указан характер преобразования координат. Например, $R_5(yzx)$ есть поворот вокруг оси C_3 (объемной диагонали Od на рис. II.8) на угол $2\pi/3$, связанный с преобразованиями: $x \longrightarrow y$,

Таблица IV.2

E	$R_1(xyz)$
$3C_4^2$	$R_2(x\bar{y}\bar{z})$, $R_3(\bar{x}y\bar{z})$, $R_4(\bar{x}\bar{y}z)$
$8C_3$	$R_5(yzx)$, $R_6(\bar{y}z\bar{x})$, $R_7(\bar{y}\bar{z}x)$, $R_8(y\bar{z}\bar{x})$, $R_9(zxy)$, $R_{10}(\bar{z}\bar{x}y)$, $R_{11}(z\bar{x}\bar{y})$, $R_{12}(\bar{z}x\bar{y})$
$6S_4=6JC_4$	$R_{13}(\bar{x}z\bar{y})$, $R_{14}(\bar{x}\bar{z}y)$, $R_{15}(\bar{z}\bar{y}x)$, $R_{16}(z\bar{y}\bar{x})$, $R_{17}(y\bar{x}\bar{z})$, $R_{18}(\bar{y}x\bar{z})$
$6\sigma=6JC_2$	$R_{19}(xzy)$, $R_{20}(x\bar{z}\bar{y})$, $R_{21}(zyx)$, $R_{22}(\bar{z}y\bar{x})$, $R_{23}(yxz)$, $R_{24}(\bar{y}\bar{x}\bar{z})$

[1]) Более подробный групповой анализ электронных спектров в кубических кристаллах см. в цитированной выше (гл. III, § 8, п. 2) статье Баукарта, Смолуховского и Вигнера.

$y \to z$, $z \to x$; $R_2(x\bar{y}\bar{z})$ — поворот вокруг оси x на угол π (C_4^2), при этом $x \to x$, $y \to -y$, $z \to -z$.

При действии преобразований R_i на функцию $f(x, y, z)$ получим, к примеру

$$\hat{R}_2\hat{R}_5 f(x, y, z) = \hat{R}_2 f(y, z, x) = f(\bar{y}, \bar{z}, x).$$

Как следует из табл. IV.2, это эквивалентно действию элемента $\hat{R}_7 f(x, y, z) = f(\bar{y}, \bar{z}, x)$, поэтому $R_2R_5 = R_7$, что может быть проверено непосредственно геометрически. Таким образом, может быть составлена вся таблица умножения элементов группы $\boldsymbol{T_d}$. Элементам S_4 соответствуют вращения на угол $\pm\pi/2$ вокруг осей x, y, z (рис. II.8, *б*) с последующим отражением в плоскости, перпендикулярной оси вращения. Легко видеть, что преобразование S_4 эквивалентно преобразованию JC_4, если вращения на угол $\pi/2$ в первом и втором случаях производить в разных направлениях; таким образом, $S_4 = JC_4$. Элементы σ связаны с отражением в плоскостях $y = \pm z$, $z = \pm x$, $x = \pm y$; они могут быть представлены как вращения на угол π с последующей инверсией (это прямо следует из (II.3.2)). Например, $R_{19}(xzy)$, связанное с отражением в плоскости $y = z$, можно представить как вращение на угол π вокруг оси $y = -z$ ($x = 0$) с последующей инверсией (рис. IV.21), поэтому $\sigma = JC_2$. Группа симметрии тетраэдра состоит из 24 элементов и пяти классов: E, $3C_4^2$, $8C_3$, $6S_4 = 6JC_4$, $6\sigma = 6JC_2$ (последовательность классов выбрана такой, как в табл. IV.2, а не такой, как в табл. II.7). Изоморфная с $\boldsymbol{T_d}$ кубическая группа $\boldsymbol{O}$ получится, если элементы двух последних классов $\boldsymbol{T_d}$ подвергнуть преобразованию инверсии; таким образом, группа $\boldsymbol{O}$ состоит из классов: E, $3C_4^2$, $8C_3$, $6C_4$, $6C_2$. Для того чтобы получить таблицу преобразований координат группы $\boldsymbol{O}$, надо в табл. IV.2 во всех двенадцати последних элементах R_i ($i = 13, 14, \ldots, 24$) произвести инверсию (тогда, например, $R_{13}(\bar{x}\bar{z}\bar{y}) \to R_{13}(xzy)$).

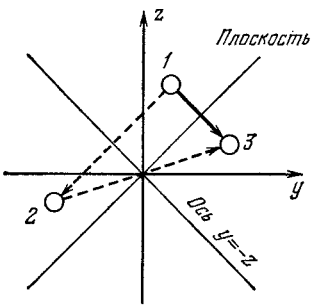


Рис. IV.21.

Полная кубическая группа $\boldsymbol{O_h}$ может быть представлена как прямое произведение группы $\boldsymbol{O}$ или $\boldsymbol{T_d}$ на $\boldsymbol{C_i} = \{E, J\}$, т. е. $\boldsymbol{O_h} = \boldsymbol{O} \times \boldsymbol{C_i} = \boldsymbol{T_d} \times \boldsymbol{C_i}$. Группа $\boldsymbol{O_h}$ состоит из 48 элементов и 10 классов, получаемых от умножения 5 классов $\boldsymbol{O}$ или $\boldsymbol{T_d}$ на элементы E и J. В первом случае ($\boldsymbol{O_h} = \boldsymbol{O} \times \boldsymbol{C_i}$) мы получим для $\boldsymbol{O_h}$ следующие 10 классов: E, $3C_4^2$, $8C_3$, $6C_4$, $6C_2$, J, $3JC_4^2$, $8JC_3$, $6JC_4$, $6JC_2$; во втором случае, при умножении классов $\boldsymbol{T_d}$ на $\boldsymbol{C_i}$, мы

получим: E, $3C_4^2$, $8C_3$, $6S_4=6JC_4$, $6\sigma=6JC_2$, J, $3JC_4^2$, $8JC_3$, $6J^2C_4=6C_4$, $6J^2C_2=6C_2$, что отличается только порядком от результата умножения $\boldsymbol{O}\times\boldsymbol{C}_i$.

Для того чтобы получить таблицу преобразований координат x, y, z для группы $\boldsymbol{O}_h$, надо дополнить табл. IV.2.24 преобразованиями R_1', R_2', ..., R_{24}', получаемыми из нее действием инверсии, так что, например, $R_{15}'=JR_{15}(\bar{z}\bar{y}x)=R_{15}'(zyx)$ (заметим, что R_{15}' принадлежит классу $6JC_4=6C_4$.

На основе такой полной таблицы для $\boldsymbol{O}_h$ можно определить, как преобразуются полиномы, содержащие слагаемые вида $x^my^nz^p$. При этом возникает возможность определить такие простейшие полиномы (с наименьшими значениями m, n, p), которые обладают симметрией базисных функций для соответствующих неприводимых представлений.

Так, например, функции $\psi=xyz$ преобразуются по неприводимому представлению Γ_2' (табл. III.2). В самом деле, используя табл. IV.2, легко показать, что, например, при действии элементов класса $3C_4^2$, $xyz\rightarrow xyz$, т. е. $\psi\rightarrow\psi$, а при действии элементов класса $6C_2$, $xyz\rightarrow -xyz$, т. е. $\psi\rightarrow -\psi$; это находится в согласии с соответствующими характерами представления Γ_2'. Легко показать, что преобразования $\boldsymbol{O}_h$ над $\psi=xyz$ полностью осуществляет представление Γ_2'.

Так же можно показать, что три функции $\psi_1=x$, $\psi_2=y$, $\psi_3=z$ преобразуются по неприводимому представлению Γ_{15}. В самом деле, для $C_4^2=R_2(\bar{x}y\bar{z})$ имеем

$$x'=1\cdot x+0\cdot y+0\cdot z,\quad y'=0\cdot x-1\cdot y+0\cdot z,$$
$$z'=0\cdot x+0\cdot y-1\cdot z.$$

Шпур (характер) этого преобразования (так же как и для двух других C_4^2) равен -1 в согласии с табл. III.2. Легко проверить, что функции x, y, z осуществляют правильные значения всех характеров для Γ_{15}.

Проверим еще, что для представления Γ_{25}' базисные функции $\psi_1=xy$, $\psi_2=xz$, $\psi_3=yz$. Например, для $C_3=R_{12}(\bar{z}x\bar{y})$ имеем

$$\hat{R}_{12}\psi_1=0\cdot\psi_1-1\cdot\psi_2+0\cdot\psi_3,\quad \hat{R}_{12}\psi_2=0\cdot\psi_1+0\cdot\psi_2+1\cdot\psi_3,$$
$$\hat{R}_{12}\psi_3=-1\cdot\psi_1+0\cdot\psi_2+0\cdot\psi_3.$$

Шпур этого преобразования (так же как для остальных C_3) равен нулю в согласии с таблицей характеров. Нетрудно показать, что приведенные выше три функции дают все характеры представления Γ_{25}'.

Действуя таким образом, мы можем составить табл. IV.3 полиномов наинизшей степени (без нормировочного множителя и при условии $x^2+y^2+z^2=r^2=\text{const}=1$), определяющих сим-

метрию базисных функций для всех неприводимых представлений группы O_h [1]).

Таблица IV.3

Γ_1	1;
Γ_2	$x^4(y^2-z^2)+y^4(z^2-x^2)+z^4(x^2-y^2)$;
Γ_{12}	$z^2-{}^1/_2(x^2+y^2)$, x^2-y^2;
Γ'_{15}	$xy(x^2-y^2)$, $yz(y^2-z^2)$, $zx(z^2-x^2)$;
Γ'_{25}	xy, yz, zx;
Γ'_1	$xyz[x^4(y^2-z^2)+y^4(z^2-x^2)+z^4(x^2-y^2)]$;
Γ'_2	xyz;
Γ'_{12}	$xyz[z^2-{}^1/_2(x^2+y^2)]$, $xyz(x^2-y^2)$;
Γ'_{15}	x, y, z;
Γ_{25}	$z(x^2-y^2)$, $x(y^2-z^2)$, $y(z^2-x^2)$.

2. В гл. I, § 2, п. 5 мы уже упоминали, что соединение сурьмянистого индия (InSb) кристаллизуется в алмазоподобной решетке, которую можно представить себе как две гранецентрированные кубические решетки (одна из атомов In, другая — из атомов Sb), сдвинутые друг относительно друга вдоль объемной диагонали куба на $^1/_4$ ее длины. InSb не обладает несобственными элементами симметрии, как германий, у которого все узлы заняты одинаковыми атомами, поэтому он описывается симморфной пространственной группой T_d^2.

[1]) У читателя может возникнуть вопрос, как находятся базисные функции в общем случае. Для определения базисных функций существует весьма громоздкий метод проекционных операторов (который мы не излагали). Р. Нокс и А. Голд (Симметрия в твердом теле.— М., 1970, с. 42) по этому поводу пишут: «У него (*т. е. читателя, А. А.*) может даже сложиться впечатление (почерпнутое из литературы), что базисные функции находят либо по догалке, либо если повезет, либо, наконец, с помощью черной магии».

Первой зоной Бриллюэна для решетки гранецентрированного куба является четырнадцатигранник, описанный в гл. IV, § 6, п. 1, изображенный на рис. IV.5 и IV.22.

Кристаллографической группой $\mathcal{F}$ решетки InSb является точечная группа тетраэдра T_d, описанная в гл. II, § 3. Эта группа состоит из 24 элементов, распределенных по пяти классам (см. табл. IV.2). Оси C_4 совпадают с осями k_x, k_y и k_z, проходящими через середины шести квадратов четырнадцатигранника. Четыре оси C_3 проходят через центры восьми шестиугольников (рис. IV.22).

Из гл. II § 9 п. 4 следует, что в случае симморфных групп ($\alpha = 0$) неприводимые представления группы волнового вектора $\Gamma(g)$ описываются неприводимыми представлениями $\Gamma(R)$, где R — элементы точечной группы $\mathcal{F}_k$. Из табл. II.7 характеров группы T_d следует, что в центре бриллюэновской зоны ($k = 0$) кристалла InSb возможны только пять состояний электрона: два невырожденных (Γ_1, Γ_2), одно дважды вырожденное (Γ_{12}) и два трижды вырожденные (Γ_{15}, Γ_{25}). Конечно, каждому из этих состояний может соответствовать множество уровней (зон) энергии, аналогично тому, как неприводимому представлению электрона в атоме (с заданным азимутальным квантовым числом l) соответствует множество уровней энергии.

Таблица IV.4

	E	C_4^2	JC_2	JC_2'
Δ_1	1	1	1	1
Δ_2	1	1	—1	—1
Δ_3	1	—1	1	—1
Δ_4	1	—1	—1	1
Γ_{15}	3	—1	—1	—1

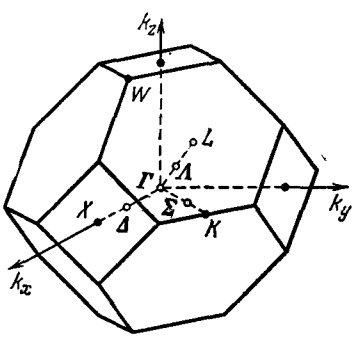


Рис. IV. 22.

Рассмотрим теперь точечную группу, соответствующую точке Δ, лежащей на оси k_x (см. рис. IV.22); она состоит из элементов R_l (см. табл. IV.2), связанных с преобразованием $x \to x$, т. е. $R_1 = E$, $R_2 = C_4^2$, $R_{19} = \sigma = JC_2$, $R_{20} = \sigma' = JC_2'$ (плоскости σ и σ' проходят через ось k_x). В табл. IV.4 даны характеры этой группы четвертого порядка, состоящей из четырех классов.

Рассмотрим точку X на поверхности зоны Бриллюэна (см. рис. IV.22). Так как волновой вектор — k_X эквивалентен век-

тору k_X, то для определения группы волнового вектора k_X надо к четырем элементам группы точки Δ добавить все элементы группы, соответствующие преобразованию $x \to -x$. Из табл. IV.2 следует, что это элементы: R_3, R_4, R_{13}, R_{14}, так что мы получаем группу восьмого порядка, состоящую из пяти классов: $R_1 = E$, $R_2 = C_4^2$, $R_3 + R_4 = 2C_4^2$, $R_{13} + R_{14} = 2JC_4 = 2S_4$, $R_{19} + R_{20} = 2JC_2 = 2\sigma$.

Таблица IV.5

	E	C_4^2	$2C_4^2$	$2JC_4$	$2JC_2$
X_1	1	1	1	1	1
X_2	1	1	1	—1	—1
X_3	1	1	—1	—1	1
X_4	1	1	—1	1	—1
X_5	2	—2	0	0	0
Γ_{15}	3	—1	—1	1	—1

В табл. IV.5 даны характеры представлений группы волнового вектора k_X. Аналогично можно построить таблицы характеров групп волнового вектора для других симметричных точек (Λ, L, Σ, K, W) зоны Бриллюэна (см. рис. IV.22).

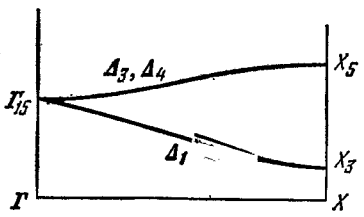


Рис. IV. 23.

Допустим, что состояние электрона проводимости в центре бриллюэновской зоны описывается трижды вырожденным неприводимым представлением Γ_{15}, тогда можно по общему правилу (гл. II, § 6, п. 4) определить, что произойдет с Γ_{15} при переходе к точкам Δ и X. Аналогично тому, как мы поступали в гл. III, § 8, при исследовании расщепления спектра колебаний получим, используя табл. IV.4 и табл. IV.5:

$$\Gamma_{15} = \Delta_1 + \Delta_3 + \Delta_4 \tag{8.1}$$

и

$$\Gamma_{15} = X_3 + X_5. \tag{8.2}$$

Дальше будет показано, что состояния Δ_3 и Δ_4 не расщепляются на оси Δ из-за дополнительной симметрии, связанной с инвариантностью уравнения Шредингера относительно обращения времени ($t \to -t$) (см. ниже § 14, п. 4), поэтому электронный спектр в InSb может иметь вид, изображенный на рис. IV.23.

Так как мы можем определить не только кратность вырождения, но и симметрию волновых функций состояний в точках Δ и X, то становится очевидной ценность информации, полученной из группового анализа для численных расчетов электронного спектра в кристаллах InSb.

§ 9. Группы волнового вектора для решетки типа германия [1])

Существенной особенностью решетки германия является наличие у нее элементов симметрии, содержащих несобственную трансляцию $\boldsymbol{\alpha} = \frac{a}{4}$ (1, 1, 1) (гл. II, § 5, п. 2); это связано с тем, что в противоположность InSb все узлы решетки заняты теперь атомами одного сорта (Ge). Бриллюэновская зона и в случае германия имеет вид четырнадцатигранника, изображенного на рис. IV.22.

Решетка германия принадлежит кристаллическому классу $\boldsymbol{O_h}$. Пространственная группа германия состоит из элементов $\{R_i | \boldsymbol{a}\}$ и $\{R_i' | \boldsymbol{\alpha} + \boldsymbol{a}\}$ ($i = 1, 2, \ldots, 24$); здесь R_i — элементы точечной группы $\boldsymbol{T_d}$ (см. табл. IV.2), $R_i' = JR_i$ (J — инверсия), $\boldsymbol{\alpha} = \frac{a}{4}(1, 1, 1)$ — несобственная трансляция и $\boldsymbol{a} \equiv \boldsymbol{a}_n$ — вектор решетки. Выберем узел решетки A на рис. I.12 за начало правой системы прямоугольных координат и направим оси x, y, z по ребрам куба (ось z вниз); основные векторы $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ направим из A к центрам прилегающих граней, т. е. к узлам *2*, *4*, *3*; в этом случае можно основные векторы решетки записать через прямоугольные составляющие в следующем виде:

$$\boldsymbol{a}_1 = \frac{a}{2}(1, 1, 0), \quad \boldsymbol{a}_2 = \frac{a}{2}(0, 1, 1), \quad \boldsymbol{a}_3 = \frac{a}{2}(1, 0, 1).$$

Хотя элементы R_i и $\{R_i' | \boldsymbol{\alpha}\}$ являются преобразованиями симметрии кристалла, они не образуют групп. Это следует из того, что произведение двух элементов такой совокупности, вообще говоря,

---

[1]) При изложении §§ 9—12 мы близко следуем книге Solid State Theory./Ed. by P. T. Landsberg.— London, 1976.

может не принадлежать той же совокупности. Например:

$$\{R_2'(\bar{x}yz)\,|\,\boldsymbol{\alpha}\}\,\{R_5'(\bar{y}\bar{z}x)\,|\,\boldsymbol{\alpha}\} =$$

$$= R_2'R_5' + \hat{R}_2'\frac{\boldsymbol{a}}{4}(1,\ 1,\ 1) + \frac{\boldsymbol{a}}{4}(1,\ 1,\ 1) =$$

$$= R_7 + \frac{\boldsymbol{a}}{4}(-1,\ 1,\ 1) + \frac{\boldsymbol{a}}{4}(1,\ 1,\ 1) = R_7 + \frac{\boldsymbol{a}}{2}(0,\ 1,\ 1), \quad (9.1)$$

где $\frac{\boldsymbol{a}}{2}(0,\ 1,\ 1) = \boldsymbol{a}_2$ — вектор решетки. (Для определения произведения $R_2'R_5' = R_7$ и действия $\hat{R}_2'\frac{\boldsymbol{a}}{4}(1,\ 1,\ 1) = \frac{\boldsymbol{a}}{4}(-1,\ 1,\ 1)$, мы воспользовались тем, что $R_i' = JR_i$, и таблицей преобразований координат (табл. IV.2). Из (9.1) действительно видно, что произведение $R_7 + \boldsymbol{a}_2$, поскольку $\boldsymbol{a}_2$ — вектор решетки, не принадлежит совокупности R_i, $\{R_i'|\boldsymbol{\alpha}\}$, т. е. последняя не образует группы.

Из (II.9.28) и (II.9.35) следует, что в центре бриллюэновской зоны ($\boldsymbol{k} = 0$) неприводимые представления группы волнового вектора $\Gamma(g) = \Gamma(\{R\,|\,\boldsymbol{\alpha} + \boldsymbol{a}\})$ совпадают с неприводимыми представлениями кристаллографической группы $\Gamma(R)$. Из табл. III.2 для характеров группы $\boldsymbol{O}_h$ следует, что в центре бриллюэновской зоны германия (и других кристаллов кубической группы) возможны только следующие состояния электрона: 4 невырожденных (Γ_1, Γ_2, Γ_1', Γ_2'), 2 дважды вырожденных (Γ_{12}, Γ_{12}') и 4 трижды вырожденных (Γ_{15}, Γ_{25}, Γ_{15}', Γ_{25}'). Симметрия волновых функций этих 10 состояний дана в табл. IV.3). И в этом случае мы, применяя теорию групп, получаем ценную информацию о возможных состояниях электронов в центре бриллюэновской зоны.

Так как точка Δ является внутренней для зоны Бриллюэна, то группа волнового вектора $\boldsymbol{k}_\Delta$ может быть описана, как это следует из (II.9.28), неприводимыми представлениями

$$\Gamma(g) = \Gamma(R) \exp i\boldsymbol{k}_\Delta \boldsymbol{\alpha}$$

(общий для всех неприводимых представлений множитель $\exp(i\boldsymbol{k}_\Delta \boldsymbol{a})$ может быть опущен). Здесь R — элементы точечной группы, соответствующие преобразованию $x \to x$; они могут быть отобраны по табл. IV.2: $R_1 = E$, $R_2 = C_4^2$, $R_{19} = JC_2$, $R_{20} = JC_2$, $R_3' = JC_4^2$, $R_4' = JC_4^3$, $R_{13}' = C_4$, $R_{14}' = C_4$. Эта группа восьмого порядка изоморфна с группами $\boldsymbol{C}_{4v}$, $\boldsymbol{D}_4$ и $\boldsymbol{D}_{2d}$. В табл. III.3 представлены характеры этой группы, состоящей из пяти классов. Для того чтобы получить таблицу характеров группы волнового вектора в точке Δ, достаточно, как это следует из (II.9.28), 3-й и 4-й столбцы табл. III.3 (связанные несобственной трансляцией), помножить на множитель $\exp(i\boldsymbol{k}_\Delta \boldsymbol{\alpha})$.

Пусть электрону в центре бриллюэновской зоны соответствует неприводимое представление Γ_{15}, т. е. он находится в трижды вырожденном p-образном состоянии. Что произойдет с ним при

переходе в точку Δ? Так как расщепление состояния Γ_{15} должно иметь место при сколь угодно малом $\boldsymbol{k}_\Delta$, то необходимо $\boldsymbol{k}_\Delta \to 0$, т. е. воспользоваться табл. III.3. Полагая

$$\Gamma_{15} = a_1\Delta_1 + a_2\Delta_2 + a_1'\Delta_1' + a_2'\Delta_2' + a_5\Delta_5, \tag{9.2}$$

получим по общему правилу (II.6.34)

$$a_1 = a_5 = 1 \tag{9.3}$$

и остальные коэффициенты в (9.2) равными нулю. Таким образом,

$$\Gamma_{15} = \Delta_1 + \Delta_5, \tag{9.4}$$

т. е. состояние Γ_{15} расщепляется в точке Δ на невырожденное состояние Δ_1 и дважды вырожденное состояние Δ_5.

Рассмотрим точку X на пересечении оси x с поверхностью зоны Бриллюэна (см. рис. IV.22), т. е. рассмотрим группу волнового вектора $\boldsymbol{k}_X = \frac{2\pi}{a}(1, 0, 0)$, где a — ребро куба прямой решетки. Поскольку группа волнового вектора $\boldsymbol{k}_X$ несимморфна, результаты, полученные в гл. II, § 9, п. 4, неприменимы.

Группа волнового вектора $\boldsymbol{k}_X$ состоит из элементов: $\{R_i \mid \boldsymbol{a}_m\}$, $\{R_i' \mid \boldsymbol{\alpha} + \boldsymbol{a}_m\}$, где $i = 1, 2, 3, 4, 13, 14, 19, 20$ (табл. IV.2), а $R_i' = JR_i$ (J — инверсия). Это симметрия точечной группы D_{4h}.

Если $\mathcal{E}_n(\boldsymbol{k})$ — d_n-кратно вырожденный уровень энергии в точке X с блоховскими функциями $\psi_{n\boldsymbol{k}_X j}(\boldsymbol{r})$ ($j = 1, 2, \ldots, d_n$), то, как следует из (II.9.12) и (II.6.1а),

$$\hat{P}_g' \psi_{n\boldsymbol{k}_X j} = \{R_i \mid \boldsymbol{a}_m\} \psi_{n\boldsymbol{k}_X j} = e^{i\boldsymbol{k}_X \boldsymbol{a}_m} \hat{R}_i \psi_{n\boldsymbol{k}_X j} =$$

$$= e^{i\boldsymbol{k}_X \boldsymbol{a}_m} \sum_{s=1}^{d_n} \Gamma(R_i)_{sj} \psi_{n\boldsymbol{k}_X s}, \tag{9.5}$$

$$\hat{P}_{g'}' \psi_{n\boldsymbol{k}_X j} = \{R_i' \mid \boldsymbol{\alpha} + \boldsymbol{a}_m\} \psi_{n\boldsymbol{k}_X j} = e^{i\boldsymbol{k}_X \boldsymbol{a}_m} \{R_i' \mid \boldsymbol{\alpha}\} \psi_{n\boldsymbol{k}_X j} =$$

$$= e^{i\boldsymbol{k}_X \boldsymbol{a}_m} \sum_{s=1}^{d_n} \Gamma(R_i')_{sj} \psi_{n\boldsymbol{k}_X s}. \tag{9.5а}$$

Заметим, что шестнадцать операций R_i и $\{R_i' \mid \boldsymbol{\alpha}\}$ ($i = 1, 2, 3, 4, 13, 14, 19, 20$) не образуют группу, как уже отмечалось выше.

Используя приведенные выше выражения основных векторов ($\boldsymbol{a}_i = 1, 2, 3$) через прямоугольные составляющие, запишем вектор решетки $\boldsymbol{a}_m = m_1\boldsymbol{a}_1 + m_2\boldsymbol{a}_2 + m_3\boldsymbol{a}_3$ через прямоугольные составляющие: $\boldsymbol{a}_m = \frac{a}{2}(m_1 + m_3,\ m_1 + m_2,\ m_2 + m_3)$. Экспоненциальный множитель, стоящий в правой части (9.5) и (9.5а), равен

$$e^{i\boldsymbol{k}_X \boldsymbol{a}_m} = e^{i\pi(m_1 + m_3)} = \begin{cases} 1, & \text{если } m_1 + m_2 \text{ — четное число,} \\ -1, & \text{если } m_1 + m_3 \text{ — нечетное число.} \end{cases} \tag{9.6}$$

Обозначая через $\boldsymbol{a}_g$ любую трансляцию, для которой $e^{i\boldsymbol{k}_X\boldsymbol{a}_g}=1$, а через $\boldsymbol{a}_n$ любую трансляцию, для которой $e^{i\boldsymbol{k}_X\boldsymbol{a}_n}=-1$, получим из (9.5) и (9.5а), что элементам группы волнового вектора $\boldsymbol{k}_X$:

$$\left.\begin{aligned} A_i &\equiv \{R_i\,|\,\boldsymbol{a}_g\}\\ B_i &\equiv \{R_i\,|\,\boldsymbol{a}_n\}\\ C_i &\equiv \{R_i'\,|\,\boldsymbol{\alpha}+\boldsymbol{a}_g\}\\ D_i &\equiv \{R_i'\,|\,\boldsymbol{\alpha}+\boldsymbol{a}_n\} \end{aligned}\right\} \text{ соответствует неприводимое представление } \left\{\begin{aligned} &\tilde{\Gamma}(R_i),\\ -&\tilde{\Gamma}(R_i),\\ &\tilde{\Gamma}(R_i'),\\ -&\tilde{\Gamma}(R_i'). \end{aligned}\right. \tag{9.7}$$

32 матрицы $\tilde{\Gamma}(R_i)$, $-\tilde{\Gamma}(R_i)$, $\tilde{\Gamma}(R_i')$, $-\tilde{\Gamma}(R_i')$ ($i=1, 2, 3, 4, 13, 14, 19, 20$)[1]) образуют неприводимое представление размерности d_n группы 32-го порядка, состоящей из элементов A_i, B_i, C_i, D_i. Легко показать, что произведение двух элементов (9.7) дает элемент той же совокупности; например:

$$A_2D_3=\{R_2\,|\,\boldsymbol{a}_g\}\{R_3'\,|\,\boldsymbol{\alpha}+\boldsymbol{a}_n\}=\{R_2R_3'\,|\,R_2\boldsymbol{\alpha}+R_2\boldsymbol{a}_n+\boldsymbol{a}_g\}=$$
$$=\{R_4'\,|\,\boldsymbol{\alpha}+\boldsymbol{a}_n\}=D_4,$$

где мы воспользовались табл. IV.2[2]).

Мы можем составить таблицу умножения этой группы и, пользуясь последней, определить ее классы. Можно показать, что группа состоит из следующих 14 классов:

$$\begin{gathered} \mathbb{C}_1=A_1,\quad \mathbb{C}_2=A_3+B_3+A_4+B_4,\quad \mathbb{C}_3=A_2,\quad \mathbb{C}_4=C_{19}+D_{20},\\ \mathbb{C}_5=C_{13}+D_{13}+C_{14}+D_{14},\quad \mathbb{C}_6=C_1+D_1\quad \mathbb{C}_7=C_3+D_3+C_4+D_4,\\ \mathbb{C}_8=C_2+D_2,\quad \mathbb{C}_9=A_{19}+A_{20},\quad \mathbb{C}_{10}=A_{13}+B_{13}+A_{14}+B_{14},\\ \mathbb{C}_{11}=B_{19}+B_{20},\quad \mathbb{C}_{12}=C_{20}+D_{19},\quad \mathbb{C}_{13}=B_2,\quad \mathbb{C}_{14}=B_1. \end{gathered} \tag{9.8}$$

Таблица характеров будет содержать 14 неприводимых представлений. Для определения таблицы характеров (табл. IV.6) могут быть использованы общие соображения, развитые в гл. II, § 6, п. 3.

Однако не все 14 неприводимых представлений могут быть использованы для интерпретации электронного спектра в германии в точке X. Необходимо, чтобы характеры этих представлений удовлетворяли условиям (9.7), а это, как мы сейчас увидим, имеет место только для четырех неприводимых представлений X_1, X_2, X_3 и X_4. Например, характеры неприводимых представлений для элементов A_i и B_i одинаковы по абсолютной

[1]) Из-за того, что мы пользуемся определением $\hat{P}_R'$ (II.6.1а), неприводимое представление образуют матрицы, транспонированные матрицам, входящим в (9.5) и (9.5а) (см. (II.6.5в)).

[2]) Заметим, что все R_i и R_i', входящие в совокупность (9.7), преобразуют $x\longrightarrow\pm x$, поэтому их применение к векторам решетки $\boldsymbol{a}_n$ и $\boldsymbol{a}_g$ не меняет их четности (так что $\boldsymbol{R}_2\boldsymbol{a}_n=\boldsymbol{a}_n$).

Таблица IV.6

	$\mathbb{C}_1$	$\mathbb{C}_2$	$\mathbb{C}_3$	$\mathbb{C}_4$	$\mathbb{C}_5$	$\mathbb{C}_6$	$\mathbb{C}_7$	$\mathbb{C}_8$	$\mathbb{C}_9$	$\mathbb{C}_{10}$	$\mathbb{C}_{11}$	$\mathbb{C}_{12}$	$\mathbb{C}_{13}$	$\mathbb{C}_{14}$
M_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
M_2	1	1	1	−1	−1	1	1	1	−1	−1	−1	−1	1	1
M_3	1	−1	1	−1	1	1	−1	1	−1	1	−1	−1	1	1
M_4	1	−1	1	1	−1	1	−1	1	1	−1	1	1	1	1
M_5	2	0	−1	0	0	2	0	−2	0	0	0	0	−2	2
M_6	1	1	1	1	1	−1	−1	−1	−1	−1	−1	1	1	1
M_7	1	1	1	−1	−1	−1	−1	−1	1	1	1	−1	1	1
M_8	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	1
M_9	1	−1	1	1	−1	−1	1	−1	−1	1	−1	1	1	1
M_{10}	2	0	−2	0	0	−2	0	2	0	0	0	0	−2	2
X_1	2	0	2	0	0	0	0	0	2	0	−2	0	−2	−2
X_2	2	0	2	0	0	0	0	0	−2	0	2	0	−2	−2
X_3	2	0	−2	2	0	0	0	0	0	0	0	−2	2	−2
X_4	2	0	−2	−2	0	0	0	0	0	0	0	2	2	−2

величине, но противоположны по знаку (9.7), т. е. $\chi(A_i) = -\chi(B_i)$. Если A_i и B_i входят в один класс, то одновременно $\chi(A_i) = \chi(B_i)$, так что $\chi(A_i) = \chi(B_i) = 0$.

Аналогичная ситуация имеет место для неприводимых представлений элементов C_i и D_i. Просматривая классы $\mathbb{C}_k$ (9.8), видно, что для «правильных» неприводимых представлений должны равняться нулю характеры классов $\mathbb{C}_2$, $\mathbb{C}_5$, $\mathbb{C}_6$, $\mathbb{C}_7$, $\mathbb{C}_8$, $\mathbb{C}_{10}$, а это имеет место только для представлений X_1, X_2, X_3 и X_4. Выделение «правильных» неприводимых представлений можно сделать и по характерам класса $\mathbb{C}_{14} = B_1$, так как

$$\chi(B_1) = \chi[\{E \mid a_n\}] = -\sum_{s=1}^{d_n} \tilde{\Gamma}(E)_{ss} = -\sum_{s=1}^{d_n} \delta_{ss} = -d_n < 0.$$

Из табл. IV.6 видно, что для $\mathbb{C}_{14}$ этому условию удовлетворяют только четыре неприводимых представления X_1, X_2, X_3 и X_4. Таким образом, в кристаллах типа германия (алмазоподобных с атомами одного сорта) в точке X существуют только дважды вырожденные состояния (см. характеры $\mathbb{C}_1 = A_1$, для X_i).

Алмазоподобная решетка сурмянистого индия (InSb), исследованная в предыдущем параграфе, описывается симморфной пространственной группой, не содержащей нетривиальных трансляций. Группа волнового вектора в точке X состоит, как мы видели, из восьми элементов $\{R_i \mid a_m\}$, где $i = 1, 2, 3, 4, 13, 14, 19, 20$. В точке X в случае InSb, как видно из табл. IV.5, возможны как невырожденные, так и дважды вырожденные состояния, в противоположность тому, что имеет место для германия.

§ 10. Спин-орбитальное взаимодействие и двойные группы

1. До сих пор при изучении поведения электрона в кристалле мы не учитывали его спина.

Однако мы знаем, что в атомах взаимодействие магнитного момента спина электронов с их орбитальным движением приводит к расщеплению и смещению уровней энергии атомов. Так, например, взаимодействие собственного магнитного момента валентного электрона с орбитальным движением вызывает у атома натрия расщепление D-линии, равное 0,002 *эв* (дублет натрия). Для более тяжелых атомов расщепление энергетических уровней больше. Для атомов рубидия оно равно 0,03 *эв*, а для ртути — 0,23 *эв*.

Спин-орбитальное смещение в спектрах электронов проводимости в полупроводниках также тем больше, чем больше атомный номер Z. Так, для InSb ($Z_{\text{In}}=49$ и $Z_{\text{Sb}}=51$) оно больше, чем для Ge ($Z_{\text{Ge}}=32$), а для последнего более существенно, чем для Si ($Z_{\text{Si}}=14$).

Из теории спина электрона Паули известно, что собственный вектор оператора спина

$$\hat{\boldsymbol{S}}=\frac{\hbar}{2}\hat{\boldsymbol{\sigma}}, \tag{10.1}$$

где двухрядные спиновые матрицы Паули $\{\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3\}=\hat{\boldsymbol{\sigma}}$ [1]), в представлении, когда спин направлен вдоль оси x_3, равны

$$\hat{\sigma}_1=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad \hat{\sigma}_2=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\quad \hat{\sigma}_3=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{10.2}$$

Мы видим, что проекция спина на ось x_3 имеет значение, равное $\pm\hbar/2=s\hbar$, где спиновая координата $s=\pm 1/2$. Легко показать, что матрицы Паули (10.2) удовлетворяют следующим соотношениям:

$$\hat{\sigma}_k^2=\boldsymbol{I},\qquad \hat{\sigma}_k\hat{\sigma}_l=-\hat{\sigma}_l\hat{\sigma}_k=i\hat{\sigma}_m, \tag{10.2a}$$

где $\boldsymbol{I}$ — единичная матрица второго ранга и индексы k, l, m пробегают значения 1, 2, 3 в циклическом порядке.

Из опыта, а также из строгой релятивистской теории Дирака следует, что со спином электрона связан магнитный момент

$$\hat{\boldsymbol{\mu}}=-\mu_0\hat{\boldsymbol{\sigma}}, \tag{10.3}$$

где

$$\mu_0=e\hbar/2mc \tag{10.3a}$$

— *магнетон Бора*.

Рассмотрим взаимодействие магнитного момента электрона μ_0 с его орбитальным движением.

---

[1]) Как обычно, индексы 1, 2, 3 соответствуют осям x, y, z.

При движении электрона в электрическом поле напряженности $\boldsymbol{E}$ на него в системе, связанной с электроном, в нерелятивистском приближении действует магнитное поле [1])

$$\boldsymbol{H}=\left[\boldsymbol{E}\times\frac{\boldsymbol{v}}{c}\right], \tag{10.4}$$

где $\boldsymbol{v}$—скорость электрона, а c—скорость света. Если электрон движется в самосогласованном периодическом потенциале кристалла $V(\boldsymbol{r})$, то

$$\boldsymbol{E}=\frac{1}{e}\nabla V(\boldsymbol{r}), \tag{10.5}$$

где e—заряд электрона.

Энергия диполя момента $\boldsymbol{\mu}$ в магнитном поле $\boldsymbol{H}$ равна $-\boldsymbol{\mu}\boldsymbol{H}$ [2]), поэтому оператор спин-орбитального взаимодействия, как это следует из (10.3)—(10.5), равен

$$\hat{\boldsymbol{\mu}}\hat{\boldsymbol{H}}=\mu_0\hat{\boldsymbol{\sigma}}\left[\frac{1}{e}\nabla V\times\frac{\hat{\boldsymbol{p}}}{mc}\right]=-\frac{i\hbar^2}{2m^2c^2}\boldsymbol{\sigma}[\nabla V\times\nabla],$$

где мы учли, что оператор скорости $\hat{\boldsymbol{v}}=(1/m)\,\hat{\boldsymbol{p}}=-(i\hbar/m)\,\nabla$. Последовательный релятивистский вывод из уравнения Дирака дает вдвое меньшее значение для спин-орбитального взаимодействия, поэтому правильный гамильтониан спин-орбитального взаимодействия равен

$$\hat{\mathcal{H}}_{so}=-\frac{i\hbar^2}{4m^2c^2}\hat{\boldsymbol{\sigma}}[\nabla V\times\nabla]. \tag{10.6}$$

Таким образом, гамильтониан электрона в кристалле, с учетом спин-орбитального взаимодействия, равен

$$\hat{\mathcal{H}}(\hat{\boldsymbol{\sigma}},\boldsymbol{r})\equiv\hat{\mathcal{H}}(\boldsymbol{r})+\hat{\mathcal{H}}_{so}, \tag{10.7}$$

где

$$\hat{\mathcal{H}}(\boldsymbol{r})=-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r}), \tag{10.7a}$$

а $\hat{\mathcal{H}}_{so}$ равен (10.6). Введем вектор

$$\hat{\boldsymbol{P}}=\frac{\hbar}{2m^2c^2}[\nabla V\times\nabla], \tag{10.8}$$

тогда гамильтониан (10.7) может быть записан в виде

$$\hat{\mathcal{H}}(\hat{\boldsymbol{\sigma}},\boldsymbol{r})=\hat{\mathcal{H}}(\boldsymbol{r})-\frac{i\hbar}{2}\hat{\boldsymbol{\sigma}}\hat{\boldsymbol{P}}=\hat{\mathcal{H}}(\boldsymbol{r})-i\hat{\boldsymbol{S}}\hat{\boldsymbol{P}}. \tag{10.9}$$

Уравнение для собственных значений энергии $\mathcal{E}$ имеет вид

$$\hat{\mathcal{H}}(\hat{\boldsymbol{\sigma}},\boldsymbol{r})\,\Psi(s,\boldsymbol{r})=\mathcal{E}\Psi(s,\boldsymbol{r}), \tag{10.10}$$

где s—спиновая координата, принимающая только значения $s=\pm1/2$, соответствующие проекции спина на ось x_3, равны

[1]) Ландау Л. Д., Лифшиц Е. М. Теория поля.—6 изд., 1973.
[2]) Тамм И. Е., § 56.

$\pm\hbar/2$. Волновая функция

$$\Psi(s, \boldsymbol{r}) = \psi_1(\boldsymbol{r})\,\nu_1(s) + \psi_2(\boldsymbol{r})\,\nu_2(s), \tag{10.11}$$

где $\nu_1(s)$ и $\nu_2(s)$ — спиновые функции Паули (спиноры), равные амплитудам вероятности того, что спин направлен по $+x_3$ или $-x_3$, поэтому

$$\nu_1(+1/2) = 1, \quad \nu_1(-1/2) = 0, \quad \nu_2(+1/2) = 0, \quad \nu_2(-1/2) = 1. \tag{10.12}$$

Спиновые функции могут быть записаны посредством матриц с одним столбцом

$$\nu_1(s) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nu_2(s) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{10.12a}$$

что эквивалентно (10.12), и соответственно полную волновую функцию (10.11) можно записать в виде

$$\Psi(s, \boldsymbol{r}) = \begin{pmatrix} \psi_1(\boldsymbol{r}) \\ \psi_2(\boldsymbol{r}) \end{pmatrix}. \tag{10.11a}$$

Координатные части $\psi_1(\boldsymbol{r})$ и $\psi_2(\boldsymbol{r})$ волновой функции $\Psi(s, \boldsymbol{r})$ (10.11) будут, вообще говоря, различными, если учитывать связь между спином электрона и его орбитальным движением.

Пользуясь (10.2) и (10.12a) и правилом умножения матриц (П.3.15), получим

$$\begin{gathered} \sigma_1\nu_1 = \nu_2, \quad \sigma_1\nu_2 = \nu_1, \quad \sigma_2\nu_1 = i\nu_2, \quad \sigma_2\nu_2 = -i\nu_1, \\ \sigma_3\nu_1 = \nu_1, \quad \sigma_3\nu_2 = -\nu_2, \\ \boldsymbol{\sigma}^2\nu_1 = (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)\,\nu_1 = 3\nu_1, \quad \boldsymbol{\sigma}^2\nu_2 = 3\nu_2. \end{gathered} \tag{10.13}$$

Из (10.12) видно, что спиновые функции ортонормированы

$$\sum_{s=\pm 1/2} \nu_i(s)\,\nu_k(s) = \delta_{ik}. \tag{10.14}$$

Нормируя полную волновую функцию (10.11) на единицу, получим, используя (10.14),

$$\sum_{s=-1/2}^{+1/2} \int \Psi^*(s, \boldsymbol{r})\,\Psi(s, \boldsymbol{r})\,d\boldsymbol{r} = \int [\psi_1^*(\boldsymbol{r})\,\psi_1(\boldsymbol{r}) + \psi_2^*(\boldsymbol{r})\,\psi_2(\boldsymbol{r})]\,d\boldsymbol{r} =$$

$$= \int [|\psi_1(\boldsymbol{r})|^2 + |\psi_2(\boldsymbol{r})|^2]\,d\boldsymbol{r} = 1. \tag{10.15}$$

Гамильтониан $\hat{\mathcal{H}}(\hat{\boldsymbol{\sigma}}, \boldsymbol{r}) = \hat{\mathcal{H}}(\boldsymbol{r}) - i\hat{\boldsymbol{S}}\hat{\boldsymbol{P}}$ (10.9) инвариантен относительно преобразований пространственной группы кристалла $\boldsymbol{G}$. В самом деле $\hat{\mathcal{H}}(\boldsymbol{r})$ (10.7а) и скалярное произведение аксиальных векторов $\hat{\boldsymbol{S}}\hat{\boldsymbol{P}}$ при таких преобразованиях не меняется (при несобственных вращениях оба вектора $\hat{\boldsymbol{S}}$ и $\hat{\boldsymbol{P}}$ меняют знак).

Таким образом, при действии элементов пространственной группы кристалла $g = \{R\,|\,\boldsymbol{\alpha} + \boldsymbol{a}\} \in \boldsymbol{G}$ на обе части уравнения (10.10) необходимо выяснить, как действует g на волновую

функцию $\Psi(s, \boldsymbol{r})$, а для этого необходимо исследовать, как ведут себя спиновые функции (спиноры) $\nu_1(s)$ и $\nu_2(s)$ при вращении координатной системы.

2. Рассмотрим вращение координатной системы вокруг оси x_3 на бесконечно малый угол $\delta\alpha$. Обозначим соответствующий оператор вращения через $R_{x_3}(\delta\alpha)$, тогда

$$R_{x_3}(\delta\alpha) f(x_1, x_2, x_3) = f(x_1 + x_2\delta\alpha, x_2 - x_1\delta\alpha, x_3) =$$
$$= f(x_1, x_2, x_3) + \frac{\partial f}{\partial x_1} x_2\delta\alpha - \frac{\partial f}{\partial x_2} x_1\delta\alpha =$$
$$= \left[1 - \delta\alpha\left(x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}\right)\right] f(x_1, x_2, x_3). \quad (10.16)$$

Введем оператор момента количества движения

$$\hat{\boldsymbol{M}} = [\boldsymbol{r}\times\hat{\boldsymbol{p}}] = \left[\boldsymbol{r}\times\frac{\hbar}{i}\nabla\right]. \quad (10.17)$$

Тогда

$$x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1} = [\boldsymbol{r}\times\nabla]_3 = \frac{i}{\hbar}M_3 \quad (10.18)$$

и

$$R_{x_3}(\delta\alpha) = 1 - \frac{i}{\hbar}\hat{M}_3\delta\alpha. \quad (10.19)$$

Вращение на конечный угол α вокруг оси x_3 можно представить себе как n последовательных поворотов на бесконечно малые углы $\delta\alpha = \alpha/n$ при $n \to \infty$; из (10.19) следует

$$R_{x_3}(\alpha) = \lim_{n\to\infty}\left(1 - \frac{i}{\hbar}\hat{M}_3\frac{\alpha}{n}\right)^n = e^{-i\alpha\hat{M}_3/\hbar}, \quad (10.20)$$

если воспользоваться для предела известным значением[1]).

Как известно, оператор в показателе экспоненты имеет смысл бесконечного ряда, получаемого при ее разложении, т. е.

$$R_{x_3}(\alpha) = e^{-i\alpha\hat{M}_3/\hbar} = 1 - \frac{i\alpha\hat{M}_3}{\hbar} + \frac{1}{2!}\left(\frac{i\alpha\hat{M}_3}{\hbar}\right)^2 - \ldots \quad (10.21)$$

Связь между оператором вращения $\hat{R}_{x_3}(\alpha)$ и оператором момента количества движения $\hat{M}_3$ (10.20) была получена для орбитального движения (10.17). Можно показать, что она имеет более общий смысл, так как связана с изотропностью пространства. Мы будем считать, что такая связь имеет место и для собственного момента количества движения или спина, т. е.

$$R_{x_i}(\alpha) = e^{-i\alpha\hat{S}_i/\hbar} = e^{\alpha\hat{\sigma}_i/2i}, \quad (10.22)$$

где $\hat{S}_i = (\hbar/2)\hat{\sigma}_i$ — составляющая спина вдоль оси x_i.

[1]) Смирнов В. И., т. I, § 38.

Применяя (10.22) к спиновой функции $\nu_1(s)$ и используя (10.13), получим

$$R_{x_3}(\alpha)\,\nu_1(s) = e^{\alpha\hat{\sigma}_3/2i}\nu_1(s) = \left[1 + \frac{\alpha\hat{\sigma}_3}{2i} + \frac{1}{2}\left(\frac{\alpha\hat{\sigma}_3}{2i}\right)^2 + \ldots\right]\nu_1(s) = e^{\alpha/2i}\,\nu_1(s). \quad (10.23)$$

Общий поворот «штрихованной» системы (x_1', x_2', x_3') относительно «нештрихованной» системы (x_1, x_2, x_3) может быть представлен тремя углами Эйлера α, β, γ (рис. IV.24). Пусть вначале «штрихованная» система совпадает с «нештрихованной»; осуществим для первой следующие три поворота: 1) вращение на положительный угол α вокруг общей оси x_3 $(x_1x_2x_3 \longrightarrow \bar{x}_1\bar{x}_2x_3)$; 2) вращение на положительный угол β вокруг оси $\bar{\bar{x}}_2$ $(\bar{x}_1\bar{x}_2x_3 \longrightarrow \bar{\bar{x}}_1\bar{\bar{x}}_2\bar{\bar{x}}_3)$ и 3) вращение на положительный угол γ вокруг $\bar{\bar{\bar{x}}}_3$ $(\bar{\bar{x}}_1\bar{\bar{x}}_2\bar{\bar{x}}_3 \longrightarrow x_1'x_2'x_3')$ [1]).

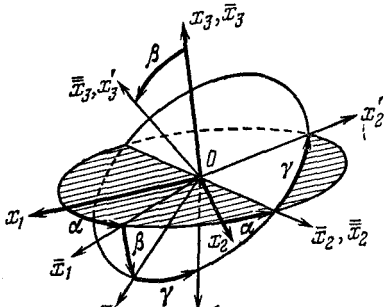


Рис. IV. 24.

Очевидно, что: $0 \leqslant \alpha \leqslant 2\pi$, $0 \leqslant \beta \leqslant \pi$, $0 \leqslant \gamma \leqslant 2\pi$. При таком повороте $(x_1x_2x_3 \longrightarrow x_1'x_2'x_2')$ прямоугольные составляющие радиуса-вектора $\boldsymbol{r} = \{x_1, x_2, x_3\}$ преобразуются следующим образом:

$$x_i' = \sum_{k=1}^{3} a_{ik}x_k. \quad (10.24)$$

Можно показать [2]), что матрица (a_{ik}) имеет следующий вид:

$$(a_{ik}) = \begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma & -\sin\beta\cos\gamma \\ -\cos\alpha\cos\beta\sin\gamma - \sin\alpha\cos\gamma & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\beta\sin\gamma \\ \cos\alpha\sin\beta & \sin\alpha\sin\beta & \cos\beta \end{pmatrix}. \quad (10.24a)$$

Если заданы углы α, β, γ, легко определить элементы матрицы a_{ik}. Обратная задача, когда дана матрица (a_{ik}) и требуется определить углы α, β и γ, решается тоже без затруднений.

---

[1]) Заметим, что разные авторы по-разному определяют углы Эйлера, от чего, конечно, зависит вид матрицы (a_{ik}) (10.24а).

[2]) Дж. Мэтьюз, Р. Уокер. Математические методы физики.— М.: Атомиздат, 1972, § 15, п. 1.

Если $a_{33}=\cos\beta\neq\pm1$, то

$$\begin{aligned}
\sin\alpha&=\pm\frac{a_{32}}{\sqrt{1-a_{33}^2}}, & \cos\alpha&=\pm\frac{a_{31}}{\sqrt{1-a_{33}^2}},\\
\sin\beta&=\pm\sqrt{1-a_{33}^2}, & \cos\beta&=a_{33},\\
\sin\gamma&=\pm\frac{a_{23}}{\sqrt{1-a_{33}^2}}, & \cos\gamma&=\mp\frac{a_{13}}{\sqrt{1-a_{33}^2}},
\end{aligned}\tag{10.25}$$

где надо использовать у всех выражений или верхний или нижний знак.

Если $a_{33}=\cos\beta=1$, т. е. $\beta=0$, то

$$\cos(\alpha+\gamma)=a_{11}=a_{22},\quad \sin(\alpha+\gamma)=a_{12}=-a_{21}.\tag{10.25а}$$

Если же $a_{33}=\cos\beta=-1$, т. е. $\beta=\pi$, то

$$\cos(\alpha-\gamma)=-a_{11}=a_{22},\quad \sin(\alpha-\gamma)=-a_{12}=-a_{21}.\tag{10.25б}$$

В случае $a_{33}=\pm1$ можно определить β и $\alpha\pm\gamma$, что оказывается в этом случае достаточным для определения спиновой матрицы $\boldsymbol{D}^l$ (10.31).

Докажем, что

$$R_{\overline{\overline{x}}_3}(\gamma)\,R_{\overline{x}_2}(\beta)\,R_{x_3}(\alpha)=R_{x_3}(\alpha)\,R_{x_2}(\beta)\,R_{x_3}(\gamma).\tag{10.26}$$

В левой части этого равенства представлены вращения на углы Эйлера α, β и γ, так как они описаны выше, т. е. вокруг осей x_3, $\overline{x}_2$ и $\overline{\overline{x}}_3$. В правой части вращения на углы γ, β и α (в обратном порядке)—вокруг неподвижных осей x_3, x_2 и x_3.

Нетрудно убедиться в том, что

$$R_{x_2}(\beta)=R_{x_3}^{-1}(\alpha)\,R_{\overline{x}_2}(\beta)\,R_{x_3}(\alpha).\tag{10.27}$$

В самом деле, в правой части мы последовательно поворачиваем подвижную систему вокруг оси x_3 на угол α, затем вращаем вокруг нового положения оси $\overline{x}_2$ на угол β и, наконец, возвращаем ось $\overline{x}_2$ в исходное положение, при котором она совпадает с осью x_2. Очевидно, это эквивалентно повороту на угол β вокруг оси x_2, т. е. равенство (10.27) доказано. Умножая слева равенство (10.27) на $R_{x_3}(\alpha)$, получим

$$R_{x_3}(\alpha)\,R_{x_2}(\beta)=R_{\overline{x}_2}(\beta)\,R_{x_3}(\alpha).\tag{10.27а}$$

Легко видеть, что

$$R_{x_3}(\gamma)=R_{x_3}^{-1}(\alpha)\,R_{\overline{x}_2}^{-1}(\beta)\,R_{\overline{\overline{x}}_3}(\gamma)\,R_{\overline{x}_2}(\beta)\,R_{x_3}(\alpha).\tag{10.28}$$

В самом деле, произведение 2-го, 3-го и 4-го множителей правой части равно $R_{x_3}(\gamma)$ (доказывается так же, как (10.27)); оставшиеся после этого справа вращения, все производятся вокруг

оси x_3, откуда прямо следует (10.28). Перемножая теперь левые и правые части равенств (10.28) и (10.27а), получим соотношение (10.26), которое нам нужно было доказать.

Из (10.26) и (10.22) следует, что оператор наиболее общего вращения $R_l(\alpha, \beta, \gamma)$, действующий на спиновые функции, равен

$$R_l(\alpha, \beta, \gamma) = R_{x_3}(\alpha)\, R_{x_2}(\beta)\, R_{x_3}(\gamma) = e^{\alpha\hat{\sigma}_3/2i} e^{\beta\hat{\sigma}_2/2i} e^{\gamma\hat{\sigma}_3/2i}. \quad (10.29)$$

Применим оператор $R_l(\alpha, \beta, \gamma)$ к спиновым функциям $\nu_1(s)$ и $\nu_2(s)$. Используя (10.13) и действуя аналогично тому, как при получении (10.23), получим после некоторых вычислений

$$\begin{aligned} R_l(\alpha, \beta, \gamma)\, \nu_1(s) &= \cos\frac{\beta}{2} e^{-\frac{i}{2}(\alpha+\gamma)} \nu_1(s) + \sin\frac{\beta}{2} e^{\frac{i}{2}(\alpha-\gamma)} \nu_2(s), \\ R_l(\alpha, \beta, \gamma)\, \nu_2(s) &= -\sin\frac{\beta}{2} e^{-\frac{i}{2}(\alpha-\gamma)} \nu_1(s) + \\ &\quad + \cos\frac{\beta}{2} e^{\frac{i}{2}(\alpha+\gamma)} \nu_2(s). \end{aligned} \quad (10.30)$$

Если рассматривать $\hat{R}_l(\alpha, \beta, \gamma)$ как элемент группы вращений, действующий на базисные спиновые функции $\nu_1(s)$ и $\nu_2(s)$, то двухрядная матрица правой части (10.30) соответствует определению оператора $\hat{P}_R$ (П. 6.1а). Неприводимым представлением, соответствующим элементу $\hat{R}_l(\alpha, \beta, \gamma)$, будет в этом случае матрица $\boldsymbol{D}^l$, транспонированная матрице (10.30), т. е.

$$\boldsymbol{D}^l = \begin{pmatrix} \cos\frac{\beta}{2} e^{-\frac{i}{2}(\alpha+\gamma)} & -\sin\frac{\beta}{2} e^{-\frac{i}{2}(\alpha-\gamma)} \\ \sin\frac{\beta}{2} e^{\frac{i}{2}(\alpha-\gamma)} & \cos\frac{\beta}{2} e^{\frac{i}{2}(\alpha+\gamma)} \end{pmatrix} \quad (10.31)$$

с характером, равным

$$\mathrm{Sp}\,\{D^l\} = \chi_l = 2\cos\frac{\beta}{2}\cos\frac{\alpha+\gamma}{2}. \quad (10.31а)$$

Рассмотрим преобразование (10.30) при повороте координатной системы вокруг любой оси на угол 2π. Такое преобразование для любой функции, зависящей от координат, эквивалентно единичному элементу E. Положим в (10.31), например, $\alpha = 2\pi$, $\beta = \gamma = 0$, тогда

$$\boldsymbol{D}^1 = \begin{pmatrix} e^{-i\pi} & 0 \\ 0 & e^{i\pi} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (10.32)$$

т. е.

$$R_1(2\pi)\, \nu_1(s) = -\nu_1(s), \quad R_1(2\pi)\, \nu_2(s) = -\nu_2(s). \quad (10.33)$$

Тот же результат получится, если положить $\beta = 2\pi$, а $\alpha = \gamma = 0$ [1]).

[1]) Конечно, то же можно показать при вращении на угол 2π вокруг оси, произвольно направленной в пространстве.

Таким образом, в то время как операция $\hat{R}_1(2\pi)$ по отношению к функциям, зависящим от пространственных координат, ведет себя как единичный элемент,— действуя на спиновые функции $\nu_1(s)$ и $\nu_2(s)$, меняет их знак (10.33). Мы обозначим

$$R_1(2\pi) \equiv \bar{E}. \tag{10.34}$$

Очевидно,

$$R_1(4\pi) = \bar{E}\bar{E} = \bar{E}^2 = E,$$

где E — единичный элемент по отношению к полной волновой функции (10.11), включающий как пространственные функции $\psi_1(\boldsymbol{r})$, $\psi_2(\boldsymbol{r})$, так и спиновые функции $\nu_1(s)$, $\nu_2(s)$.

Пусть $\hat{R}_l$ — собственные вращения и им соответствуют матрицы $\boldsymbol{D}^l$ (10.31); тогда элементам $\bar{E}R_l = \bar{R}_l$ соответствуют матрицы $-\boldsymbol{D}^l$.

Конечно, с таким же основанием мы можем элементу R_l сопоставить матрицы $-\boldsymbol{D}_l$, а элементу $\bar{E}R_l = \bar{R}_l$ — матрицу $\boldsymbol{D}^l$, так как существенным является только то, что при вращении на угол 2π матрица $\boldsymbol{D}^l$ меняет знак. Таким образом, более последовательным было бы перед матрицей в правой части (10.31) поставить знаки $\pm$.

Таким образом, в той мере, в какой учитывается спин, операции R_l не те же, что $\bar{R}_l$. В то же время, если гамильтониан $\hat{\mathcal{H}}(\hat{\sigma}, \boldsymbol{r})$ инвариантен относительно вращений R_l, то он будет инвариантен относительно операций $\bar{R}_l = \bar{E}R_l$. Таким образом, пространственная группа уравнения Шредингера — Паули (10.10) содержит вдвое больше элементов, так как наряду с $\{R_l|\boldsymbol{\alpha}+\boldsymbol{a}\}$ она содержит элементы $\{\bar{R}_l|\boldsymbol{\alpha}+\boldsymbol{a}\}$. Такие группы, учитывающие при вращениях трансформационные свойства спиновых функций, получили название *двойных групп*. Двойная группа содержит вдвое бо́льшее число элементов, чем простая, однако число классов в ней не обязательно в два раза больше. Можно показать [1]), что элементы C_2 и $\bar{E}C_2$ принадлежат к одному классу, если имеется ось C_2', перпендикулярная C_2.

В матричной форме уравнения (10.30) могут быть записаны в виде

$$\boldsymbol{R}_l(\alpha, \beta, \gamma)\begin{pmatrix}\nu_1(s)\\ \nu_2(s)\end{pmatrix} = \tilde{\boldsymbol{D}}^l\begin{pmatrix}\nu_1(s)\\ \nu_2(s)\end{pmatrix}, \tag{10.35}$$

где $\tilde{\boldsymbol{D}}^l$ — матрица, транспонированная $\boldsymbol{D}_l$ (10.31).

Сопоставляя элементу $\bar{R}_l(\alpha, \beta, \gamma)$, как указывалось выше, матрицу $-\boldsymbol{D}^l$, получим

$$\bar{\boldsymbol{R}}^l(\alpha, \beta, \gamma)\begin{pmatrix}\nu_1(s)\\ \nu_2(s)\end{pmatrix} = -\tilde{\boldsymbol{D}}^l\begin{pmatrix}\nu_1(s)\\ \nu_2(s)\end{pmatrix}. \tag{10.35а}$$

[1]) Опеховский В. О кристаллографических «двойных» группах. /В сб. Р. Нокс, А. Голд. Симметрия в твердом теле.—М., 1970, с. 271.

До сих пор мы рассматривали только собственные вращения $R_l(\alpha, \beta, \gamma)$. Покажем, что если точечная группа симметрии содержит несобственные вращения, то приведенные выше соображения о двойных группах остаются в силе.

Если группа симметрии наряду со всеми элементами собственного вращения R_l содержит несобственные вращения JR_l, то она может быть представлена как прямое произведение группы собственных вращений $\{R_l\}$ на группу $\boldsymbol{C}_i=\{E, J\}$. В гл. II, § 6, п. 5 мы показали, что неприводимые представления прямого произведения двух групп равны прямому произведению их неприводимых представлений. Так как неприводимые представления группы $\boldsymbol{C}_i$ равны $+1$ и -1 (см. табл. II.3), а выбор знака матрицы $\boldsymbol{D}^l$ для простой группы, как мы видели, произволен, то двойные группы определяются так же, как выше.

Если же операция инверсии J не входит в группу (так что группу нельзя представить как прямое произведение группы собственных вращений $\{R_l\}$ на группу $\boldsymbol{C}_i=\{E, J\}$, но группа содержит несобственные вращения JR_l, то распределение элементов по классам определяется только собственными вращениями и совпадает с распределением в изоморфных группах, содержащих вместо элементов JR_l элементы R_l. Поэтому двузначные представления таких изоморфных групп совпадают.

§ 11. Двойные группы в кристаллах InSb и Ge

1. Как уже упоминалось в § 8, п. 2, классификация электронных состояний в кристалле InSb может быть проведена по неприводимым представлениям $\Gamma(R)$ соответствующих точечных групп. Как мы видели, группе волнового вектора точки Δ (см. рис. IV.22) соответствуют следующие элементы табл. IV.2: $R_1(xyz)=E$, $R_2(x\bar{y}\bar{z})=C_4^2$, $R_{19}(xzy)=JC_2$ и $R_{20}(x\bar{z}\bar{y})=JC_2'$. Этим R_i соответствуют следующие матрицы $\boldsymbol{a}\equiv(a_{ik})$ преобразования (10.24):

$$\boldsymbol{a}^1=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\quad \boldsymbol{a}^2=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix},\quad \boldsymbol{a}^{19}=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix},$$
$$\boldsymbol{a}^{20}=\begin{pmatrix}1&0&0\\0&0&-1\\0&-1&0\end{pmatrix}. \tag{11.1}$$

Определители матриц этих преобразований равны

$$|\boldsymbol{a}^1|=|\boldsymbol{a}^2|=1,\quad |\boldsymbol{a}^{19}|=|\boldsymbol{a}^{20}|=-1, \tag{11.2}$$

откуда следует, что преобразованиям R_1 и R_2 соответствуют собственные, а R_{19} и R_{20} — несобственные вращения. Матрицы пре-

образований JR_{19} и JR_{20} (где J — инверсия), равны

$$\hat{a}_i^{19}=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \hat{a}_i^{20}=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{11.3}$$

и им соответствуют собственные вращения $(|\hat{a}_i^{19}|=|\hat{a}_i^{20}|=1)$. Можно непосредственно убедиться в том, что группа R_1, R_2, R_{19}, R_{20} изоморфна группе R_1, R_2, JR_{19}, JR_{20}.

Если $\mathcal{E}_n(\boldsymbol{k}_\Delta)$ d_n-кратно вырожденный уровень энергии в точке $\boldsymbol{k}_\Delta$, с собственными функциями $\Psi_{n\boldsymbol{k}_\Delta j}(s, \boldsymbol{r})$ $(j=1, 2, \ldots, d_n)$ (10.11), зависящими от спина, то последние являются базисными функциями неприводимого представления двойной группы размерности d_n.

Из того, что элементам R_l и $\bar{R}_l=\bar{E}R_l$ соответствуют матрицы $\boldsymbol{D}^l$ разного знака (10.35) и (10.35а), следует, что элементам двойной группы волнового вектора $\boldsymbol{k}_\Delta$: $\{R_l|\boldsymbol{a}_n\}$ и $\{\bar{R}_l|\boldsymbol{a}_n\}$ $(l=1, 2, 19, 20)$ соответствуют матрицы неприводимого представления d_n-го ранга $\Gamma(R_l)$ и $\Gamma(\bar{R}_l)=-\Gamma(R_l)$. Спиновые функции $\nu_1(s)$ и $\nu_2(s)$ сами по себе являются базисными функциями двумерного неприводимого представления двойной группы, у которой элементам R_l и $\bar{R}_l$ соответствуют матрицы второго ранга $\boldsymbol{D}^l$ и $-\boldsymbol{D}^l$ (10.31). Мы можем поэтому использовать эти матрицы для того, чтобы составить таблицу умножения нашей двойной группы.

Для определения углов Эйлера α, β, γ соответствующих собственному вращению координатной системы, мы будем пользоваться матрицами $\boldsymbol{a}^1$, $\boldsymbol{a}^2$, $\hat{\boldsymbol{a}}_i^{19}$, $\hat{\boldsymbol{a}}_i^{20}$, соответствующих изоморфной группе R_1, R_2, JR_{19}, JR_{20}. Из (10.25а), (10.25б) получим

$$\begin{aligned} &\boldsymbol{a}^1\colon\ \alpha=\beta=\gamma=0, \\ &\boldsymbol{a}^2\colon\ \beta=\pi,\quad \gamma-\alpha=\pi, \\ &\hat{\boldsymbol{a}}_i^{19}\colon\ \beta=\pi/2,\quad \alpha=\gamma=3\pi/2, \\ &\hat{\boldsymbol{a}}_i^{20}\colon\ \alpha=\beta=\gamma=\pi/2. \end{aligned}$$

Используя эти значения α, β, γ, получим из (10.31) следующее двумерное представление двойной группы:

$$\begin{aligned} &\boldsymbol{D}^1=\pm\begin{pmatrix}1&0\\0&1\end{pmatrix}, \quad \boldsymbol{D}^2=\pm\begin{pmatrix}0&-i\\-i&0\end{pmatrix}, \\ &\boldsymbol{D}^{19}=\pm\frac{1}{\sqrt{2}}\begin{pmatrix}i&-1\\1&-i\end{pmatrix}, \quad \boldsymbol{D}^{20}=\pm\frac{1}{\sqrt{2}}\begin{pmatrix}-i&-1\\1&i\end{pmatrix}, \end{aligned} \tag{11.4}$$

где условно верхний знак соответствует R_i, а нижний знак — $\bar{R}_i$ [1]. Восемь матриц (11.4) представляют некоторую группу восьмого порядка. Мы можем составить таблицу умножения этой группы.

[1] Условность этого связана с тем, что прибавление 2π к любому углу Эйлера, что всегда возможно при его определении, меняет знак матрицы $\boldsymbol{D}^l$.

Например,

$$\bar{R}_{19}R_2 = -\frac{1}{\sqrt{2}}\begin{pmatrix} i & -1 \\ 1 & -i \end{pmatrix}\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = -\frac{1}{\sqrt{2}}\begin{pmatrix} -i & -1 \\ 1 & i \end{pmatrix} = R_{20},$$

и так для всех элементов. В результате мы получим табл. IV.7, в которой первый левый столбец соответствует первому множителю, а первая верхняя строка — второму множителю. Из этой таблицы непосредственно определяются элементы, обратные данному; так, например, $\bar{R}_{19}R_{19} = R_1 \equiv E$, поэтому $R_{19}^{-1} = \bar{R}_{19}$ и $\bar{R}_{19}^{-1} = R_{19}$. Группа не абелева: $R_2R_{19} = R_{20}$, а $R_{19}R_2 = \bar{R}_{20}$, так что $R_2R_{19} \neq R_{19}R_2$.

Таблица IV.7

	R_1	$\bar{R}_1$	R_2	$\bar{R}_2$	R_{19}	$\bar{R}_{19}$	R_{20}	$\bar{R}_{20}$
R_1	R_1	$\bar{R}_1$	R_2	$\bar{R}_2$	R_{19}	$\bar{R}_{19}$	R_{20}	$\bar{R}_{20}$
$\bar{R}_1$	$\bar{R}_1$	R_1	$\bar{R}_2$	R_2	$\bar{R}_{19}$	R_{19}	$\bar{R}_{20}$	R_{20}
R_2	R_2	$\bar{R}_2$	$\bar{R}_1$	R_1	R_{20}	$\bar{R}_{20}$	$\bar{R}_{19}$	R_{19}
$\bar{R}_2$	$\bar{R}_2$	R_2	R_1	$\bar{R}_1$	$\bar{R}_{20}$	R_{20}	R_{19}	$\bar{R}_{19}$
R_{19}	R_{19}	$\bar{R}_{19}$	$\bar{R}_{20}$	R_{20}	$\bar{R}_1$	R_1	R_2	$\bar{R}_2$
$\bar{R}_{19}$	$\bar{R}_{19}$	R_{19}	R_{20}	$\bar{R}_{20}$	R_1	$\bar{R}_1$	$\bar{R}_2$	R_2
R_{20}	R_{20}	$\bar{R}_{20}$	R_{19}	$\bar{R}_{19}$	$\bar{R}_2$	R_2	$\bar{R}_1$	R_1
$\bar{R}_{20}$	$\bar{R}_{20}$	R_{20}	$\bar{R}_{19}$	R_{19}	R_2	$\bar{R}_2$	R_1	$\bar{R}_1$

Пользуясь табл. IV.7 и определением сопряженных элементов (П.2.3), можно показать, что рассматриваемая двойная группа состоит из пяти классов

$$\mathbf{C}_1 = R_1, \quad \mathbf{C}_2 = \bar{R}_1, \quad \mathbf{C}_3 = R_2 + \bar{R}_2, \quad \mathbf{C}_4 = R_{19} + \bar{R}_{19},$$
$$\mathbf{C}_5 = R_{20} + \bar{R}_{20}. \tag{11.5}$$

Так как число неприводимых представлений равно числу клас-

сов (II.6.30), а сумма квадратов размерностей неприводимых представлений равна порядку группы (II.6.22), то в рассматриваемом случае

$$1^2+1^2+1^2+1^2+2^2=8,$$

т. е. размерности неприводимых представлений Δ_1, Δ_2, Δ_3, Δ_4 и Δ_5 равны соответственно 1, 1, 1, 1 и 2; эти же числа должны стоять в первом столбце $\mathbb{C}_1=R_1$ таблицы характеров.

Таблица IV.8

	$\mathbb{C}_1$	$\mathbb{C}_2$	$\mathbb{C}_3$	$\mathbb{C}_4$	$\mathbb{C}_5$
Δ_1	1	1	1	1	1
Δ_2	1	1	1	—1	—1
Δ_3	1	1	—1	1	—1
Δ_4	1	1	—1	—1	1
Δ_5	2	—2	0	0	0

Используя свойства ортонормированности строк и ортонормированности столбцов, составим табл. IV.8 характеров нашей двойной группы. Эта таблица характеров получена обычным способом. Однако для нашей двойной группы должно выполняться условие $\chi(R_i)=-\chi(\bar{R}_i)$, поэтому, если R_i и $\bar{R}_i$ входят в один класс, то его характер должен равняться нулю. Согласно (11.5) классы $\mathbb{C}_3$, $\mathbb{C}_4$ и $\mathbb{C}_5$ удовлетворяют этому условию, поэтому соответствующие им характеры должны равняться нулю. Из табл. IV.8 видно, что это имеет место только для представления Δ_5. Можно сказать, что в точке Δ в InSb имеется для двойной группы только одно (*спинорное*) неприводимое представление Δ_5.

Мы приходим, таким образом, к заключению, что в кристаллах InSb в точке Δ электрон при учете спина может находиться только в дважды вырожденном состоянии Δ_5.

Так же могут быть рассмотрены двойные группы в InSb в точках Γ и X.

Таблица IV.9

	R_1	$\bar{R}_1$	$\frac{R_2-R_4}{\bar{R}_2-\bar{R}_4}$	R_5-R_{12}	$\bar{R}_5-\bar{R}_{12}$	$R_{13}-R_{18}$	$\bar{R}_{13}-\bar{R}_{18}$	$\frac{R_{19}-R_{24}}{\bar{R}_{19}-\bar{R}_{24}}$
Γ_6	2	—2	0	1	—1	$\sqrt{2}$	$-\sqrt{2}$	0
Γ_7	2	—2	0	1	—1	$-\sqrt{2}$	$\sqrt{2}$	0
Γ_8	4	—4	0	—1	1	0	0	0
$D\times\Gamma_{15}$	6	—6	0	0	0	$-\sqrt{2}$	$\sqrt{2}$	0

В табл. IV.9 и IV.10 представлены спинорные неприводимые представления двойной группы для InSb в центре бриллюэновской зоны Γ и в точке X; из таблиц видно, что при учете спина

Таблица IV.10

	R_1	$\overline{R}_1$	R_2, $\overline{R}_2$	R_3, R_4, $\overline{R}_3$, $\overline{R}_4$	R_{13}, R_{14}	$\overline{R}_{13}$, $\overline{R}_{14}$	R_{19}, R_{20}, $\overline{R}_{19}$, $\overline{R}_{20}$
X_6	2	−2	0	0	$\sqrt{2}$	$-\sqrt{2}$	0
X_7	2	−2	0	0	$-\sqrt{2}$	$\sqrt{2}$	0

электрон в точке Γ может находиться в двух дважды вырожденных состояниях (Γ_6, Γ_7) и одном четырежды вырожденном состоянии (Γ_8), а в точке X — в двух дважды вырожденных состояниях (X_6, X_7).

Аналогично могут быть рассмотрены двойные группы в точках симметрии бриллюэновской зоны для кристаллов Ge, пространственная группа которого не симморфна.

В табл. IV.11, IV.12, IV.13 даны спинорные неприводимые представления двойных групп в точках Γ, Δ и X для кристаллов Ge. Интересно сравнить эти таблицы с соответствующими табл. III.2, III.3 и IV.6 для простых групп в Ge в тех же точках бриллюэновской зоны. Например, электрон без учета спина мо-

Таблица IV.11

	R_1	$\overline{R}_1$	R_2-R_4, $\overline{R}_2-\overline{R}_4$	R_5-R_{12}	$\overline{R}_5-\overline{R}_{12}$	$R_{13}-R_{18}$	$\overline{R}_{13}-\overline{R}_{18}$	$R_{19}-R_{24}$, $\overline{R}_{19}-\overline{R}_{24}$
Γ_6^+	2	−2	0	1	−1	$\sqrt{2}$	$-\sqrt{2}$	0
Γ_6^-	2	−2	0	1	−1	$-\sqrt{2}$	$\sqrt{2}$	0
Γ_7^+	2	−2	0	1	−1	$-\sqrt{2}$	$-\sqrt{2}$	0
Γ_7^-	2	−2	0	1	−1	$\sqrt{2}$	$\sqrt{2}$	0
Γ_8^+	4	−4	0	−1	1	0	0	0
Γ_8^-	4	−4	0	−1	1	0	0	0

	$\{R_1' \mid \alpha\}$	$\{\overline{R}_1' \mid \alpha\}$	$\{R_2'-R_4'$, $\overline{R}_2'-\overline{R}_4' \mid \alpha\}$	$\{R_5'-R_{12}' \mid \alpha\}$	$\{\overline{R}_5'-R_{12}' \mid \alpha\}$	$\{R_{13}'-R_{18}' \mid \alpha\}$	$\{\overline{R}_{13}'-\overline{R}_{18}' \mid \alpha\}$	$\{R_{19}'-R_{24}'$, $\overline{R}_{19}'-\overline{R}_{24}' \mid \alpha\}$
Γ_6^+	2	−2	0	1	−1	$\sqrt{2}$	$-\sqrt{2}$	0
Γ_6^-	−2	2	0	−1	1	$\sqrt{2}$	$-\sqrt{2}$	0
Γ_7^+	2	−2	0	1	−1	$-\sqrt{2}$	$\sqrt{2}$	0
Γ_7^-	−2	2	0	−1	1	$-\sqrt{2}$	$\sqrt{2}$	0
Γ_8^+	4	−4	0	−1	1	0	0	0
Γ_8^-	−4	4	0	1	−1	0	0	0

Таблица IV.12

	R_1	$\bar{R}_1$	R_2, $\bar{R}_2$	R_{19}, R_{20}, $\bar{R}_{19}$, $\bar{R}_{20}$	$\{R'_3, \bar{R}'_3, R'_4, \bar{R}'_4 \mid \alpha\}$	$\{R'_{13}, R'_{14} \mid \alpha\}$	$\{\bar{R}'_{13}, \bar{R}'_{14} \mid \alpha\}$
Δ_6	2	-2	0	0	0	$\sqrt{2}\, e^{ik_\Delta \alpha}$	$-\sqrt{2}\, e^{ik_\Delta \alpha}$
Δ_7	2	-2	0	0	0	$-\sqrt{2}\, e^{ik_\Delta \alpha}$	$\sqrt{2}\, e^{ik_\Delta \alpha}$

Таблица IV.13

	R_1	$\bar{R}_1$	R_2, $\bar{R}_2$	R_3, R_4, $\bar{R}_3$, $\bar{R}_4$	R_{13}, R_{14}	$\bar{R}_{13}$, $\bar{R}_{14}$	R_{19}, R_{20}, $\bar{R}_{19}$, $\bar{R}_{20}$	$\{R'_1 \mid \alpha\}$	$\{\bar{R}'_1 \mid \alpha\}$
X_5	4	-4	0	0	0	0	0	-4	4

	$\{R'_2, \bar{R}'_2 \mid \alpha\}$	$\{R'_3, R'_4, \bar{R}'_3, \bar{R}'_4 \mid \alpha\}$	$\{R'_{13}, R'_{14} \mid \alpha\}$	$\{\bar{R}'_{13}, \bar{R}'_{14} \mid \alpha\}$	$\{R'_{19}, \bar{R}'_{19} \mid \alpha\}$	$\{R'_{20}, \bar{R}'_{20} \mid \alpha\}$
X_5	0	0	0	0	0	0

жет существовать в точке X в четырех двукратно вырожденных состояниях X_1, X_2, X_3, X_4 (табл. IV.6); при учете спина он может находиться в той же точке, только в одном четырежды вырожденном состоянии X_5 (табл. IV.13).

Все таблицы характеров для спинорных неприводимых представлений удовлетворяют отмеченному выше условию: $\chi(R_i) = -\chi(\bar{R}_i)$, поэтому, если R_i и $\bar{R}_i$ входят в один класс, то $\chi(R_i) = \chi(\bar{R}_i) = 0$.

Условия совместности неприводимых представлений, например при переходе из точки Γ на линию Δ, можно определить, пользуясь таблицами характеров.

Например, для InSb

$$\Gamma_8 = 2\Delta_5,$$

что однозначно следует из того, что Γ_8 имеет размерность 4, а Δ_5 имеет размерность 2. Полагая для Ge

$$\Gamma_7^+ = a_1\Delta_6 + a_2\Delta_7,$$

Таблица IV.14

Γ_6	Γ_7	Γ_8	X_6	X_7
Δ_5	Δ_5	$\Delta_5\Delta_5$	Δ_5	Δ_5

а)

Γ_6^+	Γ_6^-	Γ_7^+	Γ_7^-	Γ_8^+	Γ_8^-	X_5
Δ_6	Δ_6	Δ_7	Δ_7	$\Delta_6\Delta_7$	$\Delta_6\Delta_7$	$\Delta_6\Delta_7$

б)

получим из табл. IV.11, IV.12 ($k_\Delta \to 0$)

$$a_1 = \frac{1}{16}\left[2\cdot 2+(-2)(-2)+2(\sqrt{2})(-\sqrt{2})+2(-\sqrt{2})(\sqrt{2})\right] = 0,$$

$$a_2 = \frac{1}{16}\left[2\cdot 2+(-2)(-2)+2(-\sqrt{2})(-\sqrt{2})+2(\sqrt{2})(\sqrt{2})\right] = 1,$$

т. е.

$$\Gamma_7^+ = \Delta_7.$$

Таким же образом могут быть рассмотрены и остальные условия совместности, представленные в табл. IV.14 для InSb (*а*) и Ge (*б*) [1]).

§ 12. Спин-орбитальное расщепление в кристаллах In Sb и Ge

1. Предыдущий параграф был посвящен конструированию двойных групп и определению соответствующих им спинорных неприводимых представлений в некоторых симметричных точках бриллюэновской зоны, в кристаллах In Sb и Ge. Сейчас мы рассмотрим представления в центре бриллюэновской зоны Γ для In Sb и Ge, когда базисными функциями являются произведения спиновых функций на функции Блоха. Эти представления, как мы увидим, могут быть неприводимыми и приводимыми. В последнем случае это позволит определить характер спин-орбитального расщепления в точке Γ для In Sb и Ge.

[1]) Более подробно двойные группы в In Sb и Ge рассмотрены в статьях: Эллиот Р. Спин-орбитальное взаимодействие в зонной теории; таблицы характеров для некоторых «двойных» пространственных групп./В сб. Проблемы физики полупроводников.—М.: ИЛ, 1957; Парментер Р. Свойства симметрии энергетических зон в кристаллах со структурой типа цинковой обманки./В сб. Р. Нокс, А. Голд. Симметрия в твердом теле.—М., 1970, Dresselhous G.—Phys. Rev., 1955, v. 100, p. 580.

Если состояние $\mathscr{E}_n(k)$ d_n-кратно вырождено (без учета спина), т. е. ему соответствуют d_n блоховских функций $\psi_{nkj}(\boldsymbol{r})$ ($j=1, 2, \ldots, d_n$), то при учете спина состояние будет $2d_n$-кратно вырождено и волновыми функциями нулевого приближения будут $2d_n$ произведений:

$$\psi_{nkj}(\boldsymbol{r})\,\nu_i(s). \tag{12.1}$$

Блоховские волновые функции $\psi_{nkj}(\boldsymbol{r})$ удовлетворяют уравнению Шредингера

$$\hat{\mathscr{H}}(\boldsymbol{r})\,\psi_{nkj}(r)=\mathscr{E}_n(\boldsymbol{k})\,\psi_{nkj}(\boldsymbol{r}), \tag{12.2}$$

где гамильтониан

$$\hat{\mathscr{H}}(\boldsymbol{r})=-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r})$$

не зависит от спиновых операторов $\hat{\sigma}_i$; поэтому мы можем волновые функции уравнения (12.2) умножить на спиновую функцию $\nu_i(s)$ так, что

$$\hat{\mathscr{H}}(\boldsymbol{r})\,[\nu_i(s)\,\psi_{nkj}(\boldsymbol{r})]=\mathscr{E}_n(\boldsymbol{k})\,[\nu_i(s)\,\psi_{nkj}(\boldsymbol{r})]. \tag{12.2а}$$

Пусть элемент пространственной группы волнового вектора $\boldsymbol{G_k}$ есть $g=\{R_l|\boldsymbol{\alpha}_l+\boldsymbol{a}_n\}$, где R_l — собственное или несобственное вращение, а $\boldsymbol{\alpha}_l$ и $\boldsymbol{a}_n$ — несобственная трансляция, соответствующая элементу R_l и вектор решетки. Подействуем оператором $\hat{P}'_g$ (II.6.1а) на обе части уравнения (12.2а). Так как $\hat{\mathscr{H}}(r)$ инвариантен относительно действия оператора $\hat{P}'_g$, то $\hat{P}'_g[\nu_i(s)\,\psi_{nkj}(\boldsymbol{r})]$, тоже собственная функция уравнения (12.2а), соответствующая тому же собственному значению $\mathscr{E}_n(k)$, поэтому (II.7.15)

$$\hat{P}'_g[\nu_i(s)\,\psi_{nkj}(\boldsymbol{r})]=\{R_l|\boldsymbol{\alpha}_l+\boldsymbol{a}_n\}\,[\nu_i(s)\,\psi_{nkj}(\boldsymbol{r})]=$$
$$=\sum_{r=1}^{2}\sum_{s=1}^{d_n}D^l_{ri}(\alpha,\ \beta,\ \gamma)\,\Gamma(g)_{sj}\psi_{nks}(\boldsymbol{r}), \tag{12.3}$$

где α, β, γ — эйлеровы углы, соответствующие вращению R_l (если вращение несобственное, то его надо заменить собственным вращением, как это указано в конце § 10); спиновая матрица $\boldsymbol{D}^l$ определяется выражением (10.31), $\Gamma(g)$ — неприводимое представление группы $\boldsymbol{G_k}$ размерности d_n, соответствующее элементу g пространственной группы волнового вектора $\boldsymbol{k}$.

Из определения (П.3.49) видно, что

$$D^l_{ri}(\alpha,\ \beta,\ \gamma)\,\Gamma(g)_{sj}=(\boldsymbol{D}^l\times\Gamma)_{rs,\ ij} \tag{12.4}$$

— матричный элемент прямого произведения двухрядной матрицы $\boldsymbol{D}^l$ и d_n-рядной матрицы $\Gamma(g)$.

Из (12.3) и определения оператора $\hat{P}'_g$ (II.6.1а) и (II.6.5в) следует, что матричным представлением группы волнового вектора $g\in\boldsymbol{G_k}$, построенным на базисных функциях (12.1), является

транспонированная матрица (12.4). Характер этого представления

$$\chi[\boldsymbol{D}^l \times \Gamma(g)] = \chi[\boldsymbol{D}^l \times \Gamma(g)] = \chi(\boldsymbol{D}^l) \cdot \chi[\Gamma(g)] = \\ = 2\cos\frac{\beta}{2}\cos\frac{\alpha+\gamma}{2}\chi[\Gamma(g)], \quad (12.5)$$

где мы воспользовались тем, что шпур прямого произведения матриц равен произведению их шпуров (П.3.52) и выражением (10.31а). Так как матрице $\boldsymbol{D}^l$ (10.31) могут быть приписаны два знака, то разным знакам характеров (12.5) могут быть сопоставлены элементы двойной точечной группы R_l и $\bar{R}_l$.

Рассмотрим центр Г бриллюэновской зоны в кристалле InSb. Так как в точке Г волновой вектор $\boldsymbol{k}=0$, то состояние электрона в ней без учета спина может быть классифицировано по непроводимым представлениям точечной группы $\boldsymbol{T}_d$ (см. табл. II.7).

Для определения характеров прямого произведения (12.5) необходимо определить характеры матрицы $\boldsymbol{D}^l$, соответствующие разным элементам R_l группы $\boldsymbol{T}_d$. Воспользовавшись табл. IV.2 и формулами (10.25а), (10.25б), определим соответствующие углы Эйлера (они окажутся одинаковыми для всех элементов одного класса $\boldsymbol{T}_d$); после этого по формуле (10.31а) определим характеры матрицы $\boldsymbol{D}^l$.

Учитывая оба знака в (10.31а) и то, что $\chi(\bar{R}_l) = -\chi(R_l)$, можно определить характеры прямого произведения (12.5) для всех классов двойной группы $\boldsymbol{T}_d$.

Если представление двойных групп $[\boldsymbol{D}^l \times \widetilde{\Gamma(g)}]$ приводимо, то это имеет важные физические последствия. Пусть для центра бриллюэновской зоны в антимониде индия $\Gamma(g) = \Gamma(\{R \mid \boldsymbol{a}_n\}) = \Gamma(R) = \Gamma_{15}$, т. е. реализуется для простой группы неприводимое представление p-образного типа размерности 3. В этом случае прямое произведение $\boldsymbol{D} \times \Gamma_{15}$ — это матрица 6×6. Из табл. IV.9 видно, однако, что спинорные неприводимые представления двойной группы в точке Г в InSb имеют максимальную размерность 4, поэтому представление $\boldsymbol{D} \times \Gamma_{15}$ приводимо. Пользуясь формулой (12.5), определим характеры прямого произведения $\boldsymbol{D} \times \Gamma_{15}$ для всех классов двойной группы $\boldsymbol{T}_d$; они приведены в последней строке табл. IV.9, из которой мы сразу видим, что характеры $\boldsymbol{D} \times \Gamma_{15}$ равны сумме характеров Γ_7 и Γ_8, т. е.

$$\boldsymbol{D} \times \Gamma_{15} = \Gamma_7 + \Gamma_8. \quad (12.6)$$

Таким образом, шестикратно вырожденное (с учетом спина) состояние в центре бриллюэновской зоны InSb состоит из двукратно вырожденного неприводимого спинорного представления Γ_7 и четырехкратно вырожденного неприводимого представления Γ_8. Спин-орбитальное взаимодействие снимает это «случайное» вырождение и уровень распадается на два, так что $\mathscr{E}(\Gamma_8) - \mathscr{E}(\Gamma_7) =$

$=\Delta$, где Δ — спин-орбитальное расщепление. Результат (12.6) может быть интерпретирован с точки зрения электронов сильной связи. Атомный электрон в p-состоянии (азимутальное квантовое число $l=1$) характеризуется квантовым числом j, определяющим его полный (с учетом спина) момент количества движения, который равен $j\hbar$. Для $l=1$ возможные значения j равны 3/2 и 1/2; вырождение этих уровней, равное $2j+1$, принимает соответственно значения 4 и 2 аналогично (12.6). Под действием спин-орбитального взаимодействия эти уровни в атоме расщепляются. Исследование спектра атома показывает, что четырежды вырожденный терм $P_{3/2}$ расположен выше дважды вырожденного терма $P_{1/2}$. Если эта ситуация качественно сохраняется в кристалле, то $\mathscr{E}(\Gamma_8)-\mathscr{E}(\Gamma_7)=\Delta>0$.

Составляя все прямые произведения $\boldsymbol{D}^l\times\Gamma_i(R_l)$, где Γ_i неприводимые представления группы $\boldsymbol{T}_d$, и разлагая их аналогично (12.6) на спинорные неприводимые представления Γ_6, Γ_7, Γ_8, получим табл. IV.15. Мы считаем, что при $k=0$ состояние электрона в InSb в валентной зоне и зоне проводимости описывается (без учета спина) неприводимыми представлениями Γ_{15} и Γ_1.

Таблица IV.15

$\boldsymbol{D}\times\Gamma_1=\Gamma_6$, $\boldsymbol{D}\times\Gamma_2=\Gamma_7$, $\boldsymbol{D}\times\Gamma_{12}=\Gamma_8$,
$\boldsymbol{D}\times\Gamma_{15}=\Gamma_7+\Gamma_8$, $\boldsymbol{D}\times\Gamma_{25}=\Gamma_6+\Gamma_8$

На рис. IV.25 схематически изображены для InSb энергии электрона в центре бриллюэновской зоны и вдоль оси Δ, с учетом спин-орбитального взаимодействия ($б$) и без него ($а$); при этом мы использовали (8.1) и табл. IV.14 совместности.

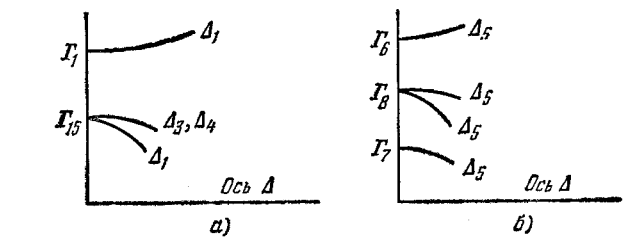


Рис. IV. 25.

На рис. IV.25, $б$ наглядно представлена величина спин-орбитального расщепления.

Аналогично табл. IV.15 может быть построена табл. IV.16 разложений прямого произведения $\boldsymbol{D}\times\Gamma_i$, где Γ_i — неприводимые представления простой группы в центре бриллюэновской

Таблица IV.16

$$D\times\Gamma_1=\Gamma_6^+,\quad D\times\Gamma_1'=\Gamma_6^-,\quad D\times\Gamma_2=\Gamma_7^+,\quad D\times\Gamma_2'=\Gamma_7^-,$$
$$D\times\Gamma_{12}=\Gamma_8^+,\quad D\times\Gamma_{12}'=\Gamma_8^-,\quad D\times\Gamma_{15}=\Gamma_6^-+\Gamma_8^-,$$
$$D\times\Gamma_{15}'=\Gamma_6^+\times\Gamma_8^+,\quad D\times\Gamma_{25}=\Gamma_7^-+\Gamma_8^-,\quad D\times\Gamma_{25}'=\Gamma_7^++\Gamma_8^+$$

зоны германия; спинорные неприводимые представления двойной группы для Ge даны в табл. IV.11. Мы считаем, что при $k=0$ состояние электрона в валентной зоне и зоне проводимости Ge описывается без учета спина неприводимыми представлениями Γ_{25}' и Γ_2'.

На рис. IV.26 схематически представлены для Ge энергии состояний электрона в точке $k=0$ и вдоль оси Δ с учетом

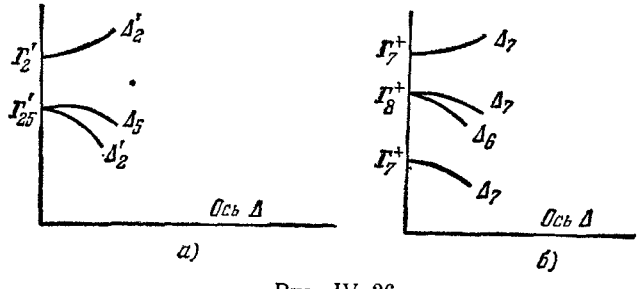


Рис. IV.26.

спин-орбитального взаимодействия (*б*) и без него (*а*); при этом мы использовали табл. III.5 и IV.14. Разность энергий состояния Γ_8^+ и нижнего состояния Γ_7^+ равна спин-орбитальному расщеплению в Ge в точке $k=0$.

§ 13. Исследование спектра электронов (дырок) вблизи минимума (максимума) энергии в зоне Бриллюэна (kp-метод)

1. В § 5 мы рассмотрели приближение почти свободных электронов, когда возмущением является периодический потенциал решетки, а в § 7 — приближение сильной связи, когда невозмущенным движением электронов является их состояние в невзаимодействующих атомах. Очевидно, что для реальных кристаллов оба приближения неудовлетворительны, поэтому существенное значение имели только качественные результаты, которые вытекали из условий пространственной периодичности решетки (вид волновой функции Блоха, существование зон разрешенной и запрещенной энергии, периодическая зависимость энергии от волнового вектора и др.).

В §§ 8—12 мы, пользуясь теорией групп, исследовали спектр электронов в симметричных точках зоны Бриллюэна. Мы смогли определить для этих точек кратность вырождения и симметрию волновых функций.

Задача определения спектра электрона $\varepsilon_n(\boldsymbol{k})$ во всей бриллюэновской зоне требует интегрирования уравнения Шредингера (3.7). Эта задача весьма сложна даже при применении электронно-вычислительных машин в большой мере из-за трудности задания самосогласованного периодического потенциала $V(\boldsymbol{r})$.

Однако для полупроводников, у которых обычно число носителей тока (электронов и дырок) мало́, достаточно определить их энергетический спектр вблизи минимумов и максимумов энергии в бриллюэновской зоне; в самом деле, если число электронов (дырок) мало, то при не слишком высоких температурах свободные носители сосредоточены в зоне вблизи минимумов (максимумов) энергии. Для определения спектра свободных носителей вблизи экстремумов энергии в бриллюэновской зоне можно применить теорию возмущений. Мы рассмотрим этот вопрос, не учитывая вначале спин-орбитального взаимодействия. Уравнение (3.8) для модулирующей функции $u_{n\boldsymbol{k}}(\boldsymbol{r})$ имеет вид

$$-\frac{\hbar^2}{2m}\nabla^2 u_{n\boldsymbol{k}} + V(\boldsymbol{r})\,u_{n\boldsymbol{k}} + \frac{\hbar}{m}(\boldsymbol{k}\hat{\boldsymbol{p}})\,u_{n\boldsymbol{k}} + \frac{\hbar^2 k^2}{2m}u_{n\boldsymbol{k}} = \varepsilon_n(\boldsymbol{k})\,u_{n\boldsymbol{k}}. \qquad (13.1)$$

Здесь n — номер разрешенной зоны энергии, поэтому $\varepsilon_n(\boldsymbol{k})$ — энергия электрона в n-й зоне в точке $\boldsymbol{k}$; $\hat{\boldsymbol{p}} = -i\hbar\nabla$ — оператор импульса электрона, m — масса свободного электрона. В точке $\boldsymbol{k}=0$ уравнение (13.1) приобретает вид

$$-\frac{\hbar^2}{2m}\nabla^2 u_{n0} + V(\boldsymbol{r})\,u_{n0} = \varepsilon_n(0)\,u_{n0}. \qquad (13.1а)$$

Для малых значений $\boldsymbol{k}$ члены

$$\frac{\hbar^2 k^2}{2m} + \frac{\hbar}{m}(\boldsymbol{kp}) \qquad (13.2)$$

в уравнении (13.1) можно рассматривать как возмущение.

Пусть экстремум энергии в зоне l расположен в точке $\boldsymbol{k}=0$, т. е. в центре бриллюэновской зоны, и пусть состояние электрона (дырки) в этой точке с энергией $\varepsilon_l(0) \equiv \varepsilon_l^0$ не вырождено. Вычислим из уравнения (13.1) поправку к энергии $\varepsilon_l(\boldsymbol{k}) - \varepsilon_l^0$ при малых значениях $\boldsymbol{k}$. Если ε_l^0 соответствует экстремуму энергии, то линейная поправка к энергии по $\boldsymbol{k}$ равна нулю. Квадратичная по $\boldsymbol{k}$ добавка к энергии состоит из поправки к энергии от $\hbar^2k^2/2m$ в первом приближении теории возмущений и от $(\hbar/m)(\hat{\boldsymbol{kp}})$ во втором приближении теории возмущений. Невозмущенные волновые функции $u_{l0}(\boldsymbol{r}) \equiv u_l$ удовлетворяют уравнению (13.1а).

Используя общие формулы теории возмущений для невырожденного состояния [1]), получим

$$\varepsilon_l(\boldsymbol{k})-\varepsilon_l^0=\left\langle u_l\left|\frac{\hbar^2k^2}{2m}\right|u_l\right\rangle+\frac{\hbar^2}{m^2}\sum_n{}'\frac{\langle u_l|\,\boldsymbol{k}\hat{\boldsymbol{p}}\,|u_n\rangle\langle u_n|\,\boldsymbol{k}\hat{\boldsymbol{p}}\,|u_l\rangle}{\varepsilon_l^0-\varepsilon_n^0}=$$

$$=\frac{\hbar^2k^2}{2m}+\frac{\hbar^2}{m^2}\sum_{\alpha\beta}k_\alpha k_\beta\sum_n{}'\frac{\langle u_l|\,\hat{p}_\alpha\,|u_n\rangle\langle u_n|\,\hat{p}_\beta\,|u_l\rangle}{\varepsilon_l^0-\varepsilon_n^0}, \qquad (13.3)$$

где штрих у суммы означает, что при суммировании опускается слагаемое с $n=l$, α и β пробегают значения x, y, z, так что, например, $\hat{p}_\alpha=\hat{p}_x=-i\hbar\partial/\partial x$ при $\alpha=x$. Матричный элемент

$$\left\langle u_l\left|\frac{\hbar^2k^2}{2m}\right|u_l\right\rangle=\frac{\hbar^2k^2}{2m}\int u_l^*u_l\,d\tau=\frac{\hbar^2k^2}{2m}=\frac{\hbar^2}{2m}\sum_{\alpha\beta}k_\alpha k_\beta\delta_{\alpha\beta}, \qquad (13.3а)$$

так как функции u_l предполагается ортонормированными. Полагая

$$\varepsilon_l(k)-\varepsilon_l^0=\frac{\hbar^2}{2}\sum_{\alpha\beta}m_{\alpha\beta}^{-1}k_\alpha k_\beta, \qquad (13.4)$$

где $m_{\alpha\beta}^{-1}$ — тензор обратной эффективной массы, получим из (13.3) и (13.4)

$$m_{\alpha\beta}^{-1}=\frac{\delta_{\alpha\beta}}{m}+\frac{2}{m^2}\sum_n{}'\frac{\langle u_l|\,\hat{p}_\alpha\,|u_n\rangle\langle u_n|\,\hat{p}_\beta\,|u_l\rangle}{\varepsilon_l^0-\varepsilon_n^0}, \qquad (13.5)$$

где $\delta_{\alpha\beta}$ — символ Кроникера. Интересно сравнить (13.5) с выражением (3.19) для той же величины $m_{\alpha\beta}^{-1}$.

В кубическом кристалле для скалярной эффективной массы получим

$$\frac{1}{m^*}=\frac{1}{m}+\frac{2}{m^2}\sum_n{}'\frac{|\langle u_l|\,\hat{p}_\alpha\,|u_n\rangle|^2}{\varepsilon_l^0-\varepsilon_n^0}. \qquad (13.6)$$

Так как ε_n^0 может быть как больше, так и меньше ε_l^0, то эффективная масса m^* может быть как меньше, так и больше массы свободного электрона m.

Формула (13.6) применима и к случаю, когда экстремум энергии $\varepsilon_l(\boldsymbol{k})$ расположен не в центре бриллюэновской зоны, а в точке $\boldsymbol{k}_0$; необходимо только учесть, что в этом случае функции u_l обладают симметрией группы волнового вектора $\boldsymbol{G}_{\boldsymbol{k}_0}$; поэтому, например, для электронов проводимости германия, у которых $\boldsymbol{k}_0$ лежит на оси [111], функция u_l обладает не кубической, а аксиальной симметрией, что приводит к тому, что в кубическом кристалле n-Ge эффективная масса является тензором.

2. Пусть теперь состояние электрона (дырки) в l-й зоне в точке $\boldsymbol{k}=0$, где энергия имеет экстремум, вырождено. Нас

[1]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика, § 38.

опять интересуют поправки к энергии, квадратичные по $\boldsymbol{k}$. Из теории возмущений вырожденных состояний следует [1]), что в нашем случае поправки к энергии $\varepsilon^{(2)}$ от $\frac{\hbar}{m}(\hat{\boldsymbol{kp}})$ во втором приближении равны корням секулярного уравнения

$$\left| \frac{\hbar^2}{m^2} \sum_{ns} \frac{\langle l,\ r' \mid \hat{\boldsymbol{kp}} \mid n,\ s\rangle \langle n,\ s \mid \hat{\boldsymbol{kp}} \mid l, r\rangle}{\varepsilon_l^0 - \varepsilon_n^0} - \varepsilon^{(2)} \delta_{rr'} \right| = 0. \qquad (13.7)$$

Здесь $|l,\ r\rangle$ и $|l,\ r'\rangle$ — невозмущенные f-кратно вырожденные волновые функции ($r,\ r' = 1,\ 2,\ \ldots,\ f$), удовлетворяющие уравнению (13.1а), для собственного значения энергии $\varepsilon_l(0) \equiv \varepsilon_l^0$; $|n,\ s\rangle$ — волновые функции для уровня энергии ε_n^0. Суммирование ведется по всем $n \neq l$ и s. Порядок определителя секулярного уравнения (13.7) равен кратности вырождения f уровня ε_l^0.

Сумма в уравнении (13.7) может быть представлена в виде

$$\frac{\hbar^2}{m^2} \sum_{\alpha\beta} k_\alpha k_\beta \sum_{ns} \frac{\langle l,\ r' \mid \hat{p}_\alpha \mid n,\ s\rangle \langle n,\ s \mid \hat{p}_\beta \mid l,\ r\rangle}{\varepsilon_l^0 - \varepsilon_n^0} \equiv \mathcal{H}_{r'r}, \qquad (13.7а)$$

где α и β принимают значения x, y, z. Запишем теперь секулярное уравнение (13.7) в виде

$$|\mathcal{H}_{r'r} - \varepsilon^{(2)} \delta_{r'r}| = 0. \qquad (13.7б)$$

Поправки к энергии ε_l^0 первого порядка от члена $\hbar^2k^2/2m$ определяются из уравнения [2])

$$\left| \left\langle l,\ r' \left| \frac{\hbar^2 k^2}{2m} \right| l,\ r \right\rangle - \varepsilon^{(1)} \delta_{r'r} \right| = 0 \qquad (13.8)$$

или

$$\left| \left(\frac{\hbar^2 k^2}{2m} - \varepsilon^{(1)} \right) \delta_{r'r} \right| = 0. \qquad (13.8а)$$

Отсюда следуют f совпадающих корней

$$\varepsilon^{(1)} = \hbar^2 k^2 / 2m. \qquad (13.9)$$

Из (13.7) и (13.9) следует, что

$$\varepsilon_l(\boldsymbol{k}) = \varepsilon_l^0 + \varepsilon^{(1)} + \varepsilon_l^{(2)} = \varepsilon_l^0 + \frac{\hbar^2 k^2}{2m} + \varepsilon_l^{(2)}, \qquad (13.10)$$

где $\varepsilon_l^{(2)}$ — один из f корней уравнения (13.7).

3. Рассмотрим структуру секулярного уравнения (13.7б) в случае валентной зоны кремния или германия. Теоретические и экспериментальные соображения делают весьма правдоподобным, предположение о том, что состояние дырки в центре зоны ($\boldsymbol{k} = 0$) соответствует неприводимому представлению Γ_{25}' группы $\boldsymbol{O}_h$. Базисные функции этого неприводимого представления с размер-

[1]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика, § 39.
[2]) См. там же.

ностью 3 следующие (см. табл. IV.3): $X^+ = xy$, $Y^+ = yz$, $Z^+ = zx$. Нетрудно показать, что эти базисные функции при действии преобразований симметрии группы O_h ведут себя аналогично более простым базисным функциям $X = x$, $Y = y$, $Z = z$ неприводимого представления Γ_{15}, соответствующего трехмерно-вырожденному атомному p-состоянию электрона; поэтому ниже мы пользуемся этими более простыми функциями с симметрией x, y, z.

Можно показать, что сумма $\sum_{ns}$ (13.7а) преобразуется при действии элементов пространственной группы кристалла, как матричный элемент

$$\langle l,\ r' \mid \hat{p}_\alpha \hat{p}_\beta \mid l,\ r\rangle \equiv Q_{\alpha\beta}(r',\ r), \tag{13.11}$$

где операторы $\hat{p}_x$, $\hat{p}_y$, $\hat{p}_z$ преобразуются как координаты x, y, z (в самом деле, если при данном преобразовании $x \to -y$, то $\hat{p}_x = -i\hbar\partial/\partial x \to i\hbar\partial/\partial y = -\hat{p}_y$).

Рассмотрим константу $Q_{\alpha\beta}(r',\ r)$ для состояний $r' = r = X$ и покажем, что для $\alpha \neq \beta$ она равна нулю; положим, например, $p_\alpha = p_x$, $p_\beta = p_y$, тогда (13.11) преобразуются под действием элементов группы O_h как произведение $[xxyx]$; при применении преобразования $R_2(x\ \bar{y}\ \bar{z})$ (табл. IV.2) получим

$$\hat{R}_2[x\,x\,y\,x] = -[x\,x\,y\,x],$$

откуда следует, что $[x\,x\,y\,x] = 0$, т. е. соответствующая константа $Q_{xy}(X,\ X)$ равна нулю. Если для тех же состояний $r' = r = X$ положить $p_\alpha = p_\beta = p_x$, то (13.11) преобразуется как произведение $[x\,x\,x\,x] = [x^4]$; легко видеть, что не может быть такого преобразования R_i или JR_i, которое преобразовывало бы $[x^4]$ в $-[x^4]$, поэтому соответствующая константа $Q_{xx}(X,\ X)$ отлична от нуля; мы обозначим ее через

$$L = \frac{\hbar^2}{m^2} \sum_{ns}{}' \frac{\langle l,\ X \mid \hat{p}_x \mid n,\ s\rangle \langle n,\ s \mid \hat{p}_x \mid l,\ X\rangle}{\varepsilon_l^0 - \varepsilon_n^0} =$$

$$= \frac{\hbar^2}{m^2} \sum_{ns}{}' \frac{|\langle l,\ X \mid \hat{p}_x \mid n,\ s\rangle|^2}{\varepsilon_l^0 - \varepsilon_n^0}. \tag{13.12}$$

Пусть по-прежнему состояния $r' = r = X$, а $p_\alpha = p_\beta = p_y$, тогда применяя $R_{13}(\bar{x}\ z\ \bar{y})$ к $[xyyx]$, получим

$$\hat{R}_{13}[xy^2x] = [xz^2x],$$

и это позволяет ввести константу

$$M = \frac{\hbar^2}{m^2} \sum_{ns}{}' \frac{|\langle l,\ X \mid \hat{p}_y \mid n,\ s\rangle|}{\varepsilon_l^0 - \varepsilon_n^0} = \frac{\hbar^2}{m^2} \sum_{ns}{}' \frac{|\langle l,\ X \mid \hat{p}_z \mid n,\ s\rangle|^2}{\varepsilon_l^0 - \varepsilon_n^0}. \tag{13.13}$$

Мы можем записать теперь (1,1)-й член определителя (13.7б) в виде

$$\mathscr{H}_{xx}-\varepsilon^{(2)}=Lk_x^2+M(k_y^2+k_z^2)-\varepsilon^{(2)}.$$

Совершенно так же можно показать, что (1,2)-й и (2,1)-й члены определителя (13.7а) равны

$$\mathscr{H}_{xy}=\mathscr{H}_{yx}=Nk_xk_y,$$

где

$$N=\frac{\hbar^2}{m^2}\sum_{ns}{}'\frac{\langle l,\ X\,|\,\hat{p}_x\,|\,n,\ s\rangle\langle n,\ s\,|\,\hat{p}_y\,|\,lY\rangle}{\varepsilon_l^0-\varepsilon_n^0}. \qquad (13.14)$$

Из табл. IV.2 следует, что $\hat{R}_5[x^4]=[y^4]$, $\hat{R}_5[xy^2x]=[yz^2y]$ и $\hat{R}_{17}[xy^2x]=[yx^2y]$, поэтому (2,2)-й член определителя (13.7б) равен

$$\mathscr{H}_{yy}-\varepsilon^{(2)}=Lk_y^2+M(k_x^2+k_z^2)-\varepsilon^{(2)}.$$

В результате секулярное уравнение (13.7б) приобретает вид

$$\begin{vmatrix} Lk_x^2+M(k_y^2+k_z^2)-\varepsilon^{(2)} & Nk_xk_y & Nk_xk_z \\ Nk_xk_y & Lk_y^2+M(k_x^2+k_z^2)-\varepsilon^{(2)} & Nk_yk_z \\ Nk_xk_z & Nk_yk_z & Lk_z^2+M(k_x^2+k_y^2)-\varepsilon^{(2)} \end{vmatrix}=0, \qquad (13.15)$$

где материальные константы L, M, N определяются выражениями (13.12), (13.13) и (13.14).

Секулярное уравнение (13.5) третьей степени относительно $\varepsilon^{(2)}$. Мы не будем решать его в общем виде, а ограничимся случаем выделенного направления волнового вектора $\boldsymbol{k}$. Пусть $k_x\neq 0$, $k_y=k_z=0$, тогда (13.5) приобретает вид

$$\begin{vmatrix} Lk_x^2-\varepsilon^{(2)} & 0 & 0 \\ 0 & Mk_x^2-\varepsilon^{(2)} & 0 \\ 0 & 0 & Mk_x^2-\varepsilon^{(2)} \end{vmatrix}=0, \qquad (13.15а)$$

откуда $[Lk_x^2-\varepsilon^{(2)}][Mk_x^2-\varepsilon^{(2)}]^2=0$.

Таким образом, мы имеем три корня уравнения (13.15а):

$$\varepsilon_1^{(2)}=Lk_x^2,\quad \varepsilon_2^{(2)}=\varepsilon_3^{(2)}=Mk_x^2, \qquad (13.15б)$$

которые полностью соответствуют случаю сильной связи (IV.7.25):

$$\varepsilon_1=\varepsilon_M-Aa^2k_x^2,\quad \varepsilon_2=\varepsilon_3=\varepsilon_M-Ba^2k_x^2.$$

Конечно, общее решение секулярного уравнения (13.15) не совпадает с простым выражением (IV.7.24), что свидетельствует о том, что $\boldsymbol{k}\hat{\boldsymbol{p}}$-метод в квадратичном приближении выходит за рамки простой модели сильной связи, учитывающей взаимодействие лишь соседних атомов.

4. Уравнение (13.15) определяет спектр дырок в кристаллах германия или кремния без учета спин-орбитального взаимодей-

ствия, которое, как мы видели в § 12, приводит к расщеплению уровней энергии.

Состоянию Γ'_{25} в центре валентной зоны при учете спина соответствуют не три функции, преобразующиеся как X, Y и Z, а шесть функций вида: $X\nu_1$, $Y\nu_1$, $Z\nu_1$, $X\nu_2$, $Y\nu_2$ и $Z\nu_2$, где $\nu_1(s)$ и $\nu_2(s)$ — спиновые функции (10.12). Секулярное уравнение (13.7) будет теперь определителем шестого порядка.

Очевидно, что в качестве волновых функций нулевого приближения можно взять шесть произвольных линейных комбинаций указанных выше шести функций $X\nu_1$, $Y\nu_1$, ..., $Z\nu_2$. Можно показать, что если выбрать эти линейные комбинации в виде

$$Y_{3/2}^{3/2}=-\frac{1}{\sqrt{2}}(x+iy)\,\nu_1,\quad Y_{3/2}^{1/2}=-\frac{1}{\sqrt{6}}[(x+iy)\,\nu_2-2z\nu_1],$$
$$Y_{3/2}^{-1/2}=\frac{1}{\sqrt{6}}[(x-iy)\,\nu_1+2z\nu_2],\quad Y_{3/2}^{-3/2}=\frac{1}{\sqrt{2}}(x-iy)\,\nu_2 \tag{13.16}$$

и

$$Y_{1/2}^{1/2}=-\frac{1}{\sqrt{3}}[(x+iy)\,\nu_2+z\nu_1],\quad Y_{1/2}^{-1/2}=\frac{1}{\sqrt{3}}[(x-iy)\,\nu_1-z\nu_2], \tag{13.16a}$$

то они диагонализуют гамильтониан спин-орбитального взаимодействия $\hat{\mathcal{H}}_{s0}$ (10.6). В результате этого секулярное уравнение с определителем шестого порядка распадается на два секулярных уравнения с определителями четвертого и второго порядков (см. ниже).

Функции Y_j^m на единичной сфере, т. е. при условии $x^2+y^2+z^2=1$ — шаровые функции, являющиеся собственными функциями оператора проекции полного момента количества движения

$$\hat{J}_z=\hat{M}_z+\hat{S}_z, \tag{13.17}$$

где $\hat{\boldsymbol{M}}$ и $\hat{\boldsymbol{S}}$ определяются выражениями (10.17) и (10.1); $m\hbar$ — проекция полного момента количества движения на ось z, максимальное значение которого равно $j\hbar$.

Применяя, например, оператор $\hat{J}_z$ (13.17) к функции $Y_{3/2}^{-1/2}$, получим

$$\hat{J}_zY_{3/2}^{-1/2}=\frac{1}{\sqrt{6}}\left[\frac{\hbar}{i}\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)+\frac{\hbar}{2}\hat{\sigma}_z\right][(x-iy)\,\nu_1+2z\nu_2]=$$
$$=-\frac{\hbar}{2}\frac{1}{\sqrt{6}}[(x-iy)\,\nu_1+2z\nu_2]=-\frac{\hbar}{2}Y_{3/2}^{-1/2},$$

где мы воспользовались тем, что $\hat{\sigma}_z\nu_1=\nu_1$ и $\hat{\sigma}_z\nu_2=-\nu_2$ (см. (10.13)). Мы видим, что собственное значение оператора $\hat{J}_z$, соответствующее функции $Y_j^m=Y_{3/2}^{-1/2}$, действительно равно $m\hbar=-\hbar/2$, т. е. $m=-1/2$. Можно показать, что волновые функции

(13.16а) на единичной сфере взаимно ортогональны и нормированы (на $4\pi/3$).

Структура «правильных» волновых функций нулевого приближения (13.16а) может быть понята и из следующих соображений. Представление двойной группы Γ_{15} равно прямому произведению $D^{1/2}\times D^1$, где D^1 — неприводимое представление Γ_{15}, а $D^{1/2}$ — неприводимое спинорное представление (10.31). Можно показать, что

$$D^{1/2}\times D^1 = D^{3/2} + D^{1/2},$$

где $D^{3/2}$ — неприводимое представление размерности 4, соответствующее полному моменту импульса, равно $^3/_2\hbar$, а $D^{1/2}$ — неприводимое спинорное представление размерности 2 с моментом $^1/_2\hbar$. Эти представления также неприводимые в кубической группе и переходят соответственно в представления Γ_8 и Γ_7 в согласии с разложением (12.6): $D^{1/2}\times\Gamma_{15} = \Gamma_8 + \Gamma_7$. Поэтому базисными функциями четырехкратно вырожденного представления Γ_8 являются функции (13.16) — собственные функции оператора полного момента J_z с максимальным собственным значением $^3/_2\hbar$, а базисными функциями двукратно вырожденного состояния Γ_7 — функции (13.16а), соответствующие моменту $^1/_2\hbar$.

Можно показать, что недиагональные матричные элементы оператора спин-орбитального взаимодействия $\hat{\mathcal{H}}_{s0}$ (10.6), построенные на функциях (13.16) и (13.16а), равны нулю. Добавка к диагональным матричным элементам $\mathcal{H}_{r'r}$ (13.7а), связанная со спин-орбитальным взаимодействием, на функциях (13.16), равна $^1/_3\Delta$, а на функциях (13.16а) равна $-^2/_3\Delta$, где Δ — спин-орбитальное расщепление зон Γ_8 и Γ_7 при $\boldsymbol{k}=0$ [1]).

Если расщепление зон $\mathscr{E}(\Gamma_8) - \mathscr{E}(\Gamma_7) = \Delta$, обусловленное спин-орбитальным взаимодействием, велико по сравнению с кинетической энергией носителей тока, то секулярные уравнения для зон Γ_8 и Γ_7 можно рассматривать раздельно.

Обозначим через $\mathcal{H}'_{ik}$ элементы матрицы секулярного уравнения для определения энергии возмущения во втором приближении при учете спин-орбитального взаимодействия, когда в качестве «правильных» волновых функций нулевого приближения выбраны функции (13.16); сопоставим у $\mathcal{H}'_{ik}$ индексам i, $k=1$, 2, 3, 4 значения: $1\to{}^3/_2$, $2\to{}^1/_2$, $3\to-{}^1/_2$, $4\to-{}^3/_2$, тогда из (13.16) получим

$$\mathcal{H}'_{11} = {}^1/_2\langle(x+iy)\,v_1|\hat{\mathcal{H}}|(x+iy)\,v_1\rangle = {}^1/_2(\mathcal{H}_{xx}+\mathcal{H}_{yy}+i\mathcal{H}_{xy}-i\mathcal{H}_{yx}),$$

[1]) Воспользовавшись выражением (10.6) для $\hat{\mathcal{H}}_{s0}$, явным видом функций (13.16), (13.16а) и соображениями симметрии, нетрудно показать, что смещение уровня энергии при $\boldsymbol{k}=0$ зоны Γ_8 пропорционально $2\langle\boldsymbol{r}|\nabla N\rangle$, а зоны Γ_7 пропорционально $(-4\langle\boldsymbol{r}|\nabla N\rangle)$, так что расщепление $\Delta\backsim 6\langle\boldsymbol{r}|\nabla V\rangle$.

где $\mathcal{H}'_{rr}(r',\ r=x,\ y,\ z)$ — элементы матрицы (13.7а); воспользовавшись обозначениями (13.12), (13.13) и учитывая, что $\mathcal{H}_{xy}=\mathcal{H}_{yx}$, получим

$$\mathcal{H}'_{11}={}^1/_2(\mathcal{H}_{xx}+\mathcal{H}_{yy})={}^1/_2(L+M)(k_x^2+k_y^2)-Mk_z^2=$$
$$=Ak^2+{}^1/_2B(k^2-3k_z^2)\equiv F, \quad (13.18)$$

где

$$A=\frac{L+2M}{3} \quad \text{и} \quad B=\frac{L-M}{3}. \quad (13.18\text{а})$$

Легко видеть, что

$$\mathcal{H}'_{44}=\mathcal{H}'_{11}=F.$$

Далее

$$\mathcal{H}'_{22}={}^1/_6\langle(x+iy)\,v_2-2zv_1\,|\,\hat{\mathcal{H}}\,|\,(x+iy)\,v_2-2zv_1\rangle=$$
$$={}^1/_6(\mathcal{H}_{xx}+\mathcal{H}_{yy}+i\mathcal{H}_{xy}-i\mathcal{H}_{yx}+4\mathcal{H}_{zz}),$$

где мы использовали ортогональность спиновых функций; подставляя сюда значения элементов матрицы (13.15), получим после некоторых преобразований

$$\mathcal{H}'_{22}=Ak^2-{}^1/_2B(k^2-3k_z^2)\equiv G. \quad (13.19)$$

Легко показать, что

$$\mathcal{H}'_{33}=\mathcal{H}'_{22}=G.$$

Далее

$$\mathcal{H}'_{12}=\frac{1}{2\sqrt{3}}\langle(x+iy)\,v_1\,|\,\hat{\mathcal{H}}\,|\,(x+iy)\,v_2-2zv_1\rangle=$$
$$=-\frac{1}{\sqrt{3}}(\mathcal{H}_{xz}-i\mathcal{H}_{yz})=-Dk_z(k_x-ik_y)\equiv H, \text{ где } D=N/\sqrt{3}, \quad (13.20)$$

$$\mathcal{H}'_{13}=-\frac{\sqrt{3}}{2}[B(k_x^2-k_y^2)+iDk_xk_y]\equiv J. \quad (13.21)$$

Легко видеть, что

$$\mathcal{H}'_{14}=\mathcal{H}'_{23}=0, \quad \mathcal{H}'_{24}=J, \quad \mathcal{H}'_{34}=-H.$$

Остальные элементы матрицы $\mathcal{H}'_{ik}$ можно получить, используя ее эрмитовость:

$$\mathcal{H}'_{ik}=\mathcal{H}'^*_{ki}.$$

Секулярное уравнение для энергии $\varepsilon\equiv\varepsilon^{(2)}$ имеет вид

$$\begin{vmatrix} F-\varepsilon & H & J & 0 \\ H^* & G-\varepsilon & 0 & J \\ J^* & 0 & G-\varepsilon & -H \\ 0 & J^* & -H^* & F-\varepsilon \end{vmatrix}=0, \quad (13.22)$$

где величины F, G, H, J равны (13.18)—(13.21).

Разлагая определитель четвертого порядка (13.22) по элементам первой строки, после некоторых алгебраических преобразований получим

$$[(F-\varepsilon)(G-\varepsilon)-|H|^2-|J|^2]^2=0,$$

или

$$(\varepsilon - F)(\varepsilon - G) - |H|^2 - |J|^2 = 0.$$

Решая это квадратное уравнение, получим

$$\varepsilon_{1,2} = \frac{F+G}{2} \pm \sqrt{\left(\frac{F+G}{2}\right)^2 - FG + |H|^2 + |J|^2}.$$

Подставляя значения F, G, H и J, получим после длинных, но простых алгебраических преобразований

$$\varepsilon_{1,2} = Ak^2 \pm \sqrt{B^2k^4 + C^2(k_x^2k_y^2 + k_x^2k_z^2 + k_y^2k_z^2)}, \qquad (13.23)$$

где A и B равны (13.18а), а

$$C^2 = D^2 - 3B^2 = \frac{(N+L-M)(N-L+M)}{3}, \qquad (13.23а)$$

если использовать значение B и D (13.20).

Так как исходное уравнение для ε биквадратно, то каждый из корней (13.23) двукратно вырожден, что связано с инвариантностью относительно инверсии времени.

Для определения спектра энергии в отщепившейся в результате спин-орбитального взаимодействия двукратно вырожденной зоне Γ_7 необходимо решить секулярное уравнение ранга 2×2, матричные элементы которого $\mathscr{H}'_{ik}$ вычисляются на волновых функциях (13.16а). Сопоставляя у $\mathscr{H}'_{ik}$ индексы i, $k = 1, 2$ значениям: $1 \longrightarrow {}^1/_2$, $2 \longrightarrow -{}^1/_2$, получим из (13.16а)

$$\mathscr{H}'_{11} = {}^1/_3 \langle (x+iy)\nu_2 + z\nu_1 | \tilde{\mathscr{H}} | (x+iy)\nu_2 + z\nu_1 \rangle =$$
$$= {}^1/_3 [\mathscr{H}_{xx} + \mathscr{H}_{yy} - i\mathscr{H}_{yx} + i\mathscr{H}_{xy} + \mathscr{H}_{zz}] =$$
$$= \frac{L+2M}{3}(k_x^2 + k_y^2 + k_z^2) = Ak^2, \qquad (13.24)$$

где мы воспользовались обозначениями (13.12), (13.13) и (13.18а).

Легко показать, что

$$\mathscr{H}'_{22} = \mathscr{H}'_{11}, \quad \mathscr{H}'_{12} = \mathscr{H}'_{21} = 0. \qquad (13.24а)$$

Используя полученные значения $\mathscr{H}'_{ik}$, составим секулярное уравнение

$$\begin{vmatrix} Ak^2 - \Delta - \varepsilon & 0 \\ 0 & Ak^2 - \Delta - \varepsilon \end{vmatrix} = 0. \qquad (13.25)$$

Здесь $\Delta = \mathscr{E}(\Gamma_8) - \mathscr{E}(\Gamma_7)$ — спин-орбитальное расщепление уровней Γ_8 и Γ_7 в точке $\boldsymbol{k} = 0$.

Из (13.25) имеем для дважды вырожденного корня

$$\varepsilon_3 = -\Delta + Ak^2, \qquad (13.26)$$

т. е. в зоне Γ_7 имеет место простой параболический закон дисперсии (поверхности постоянной энергии — сферы). Закону дисперсии энергии (13.23) для зоны Γ_8 соответствуют поверхности

постоянной энергии — *гофрированные поверхности*, имеющие в плоскости [100] (т. е. плоскости $k_x k_y$) вид, изображенный на рис. IV.27.

5. Опыты по циклотронному резонансу, явлениям переноса и оптическому поглощению в антимониде индия позволили установить следующие особенности его зонной структуры. Минимум зоны проводимости и максимум валентной зоны расположены в точке $\boldsymbol{k}=0$. Эффективная масса электронов на дне зоны проводимости очень мала и равна 0,013 m, т. е. порядка 0,01 массы свободного электрона. Две ветви дырок, тяжелых и легких, вырождены в точке $\boldsymbol{k}=0$. Эффективная масса тяжелых дырок порядка $0{,}18m$, т. е. более чем в 10 раз больше массы электронов и, по-видимому, легких дырок. Третья ветвь дырок (также легких) отщеплена за счет спин-орбитального взаимодействия от верхнего края валентной зоны на величину $\Delta \approx 0{,}9$ *эв*.

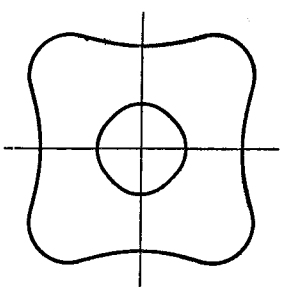


Рис. IV. 27.

Такая структура валентной зоны в антимониде индия следует из теории, если использовать симметрию кристалла InSb и не учитывать поправок к энергии линейных по k. Закон дисперсии энергии $\varepsilon(\boldsymbol{k})$ для всех ветвей обладает сферической симметрией, но для легких частиц при больших значениях k обнаруживает заметное отклонение от простой параболичности (ε не пропорционально k^2).

Ширина запрещенной зоны ε_G при комнатной температуре равна 0,17 *эв* и сильно зависит от температуры, так что при 0°K $\varepsilon_G = 0{,}23$ *эв*. Узость запрещенной зоны требует одновременного рассмотрения состояния электрона в валентной зоне и зоне проводимости.

Такое «взаимодействие» состояний электрона в валентной зоне и зоне проводимости с одновременным учетом спин-орбитального взаимодействия позволило полуколичественно объяснить большинство особенностей зонной структуры InSb (Е. О. Кейн, 1956).

Оператор спин-орбитального взаимодействия (10.6), действуя на блоховскую функцию, дает два слагаемых

$$\hat{\mathcal{H}}_{so}[u_k(\boldsymbol{r})\,e^{i\boldsymbol{k}\boldsymbol{r}}] = \\ = e^{i\boldsymbol{k}\boldsymbol{r}}\frac{\hbar}{4m^2c^2}\hat{\boldsymbol{\sigma}}[\nabla V\times\hat{\boldsymbol{p}}]\,u_k(\boldsymbol{r}) + e^{i\boldsymbol{k}\boldsymbol{r}}\frac{\hbar^2}{4m^2c^2}\hat{\boldsymbol{\sigma}}[\nabla V\times\boldsymbol{k}]\,u_k(\boldsymbol{r}). \quad (13.27)$$

По порядку величины первое слагаемое в $p/\hbar k$ раз больше второго; здесь p — импульс электрона в атоме (спин-орбитальное взаимодействие в основном обусловлено электронами атомов

решетки), а $\hbar k$ — квазиимпульс электрона проводимости; отношение $p/\hbar k$ велико.

Первое слагаемое в (13.27) определяет спин-орбитальное расщепление уровней, второе — ответственно за зависимость эффективной массы от энергии.

Удерживая только первое слагаемое, получим вместо (13.1) (опуская номер зоны n)

$$\left\{\frac{\hat{p}}{2m}+V(\boldsymbol{r})+\frac{\hbar}{m}(\hat{\boldsymbol{k}}\hat{\boldsymbol{p}})+\frac{\hbar}{4m^2c^2}\hat{\boldsymbol{\sigma}}[\nabla V\times\hat{\boldsymbol{p}}]\right\}u_{\boldsymbol{k}}(\boldsymbol{r})=\varepsilon'_{\boldsymbol{k}}u_{\boldsymbol{k}}(\boldsymbol{r}), \quad (13.28)$$

где

$$\varepsilon'_{\boldsymbol{k}}=\varepsilon_{\boldsymbol{k}}-\frac{\hbar^2k^2}{2m}. \quad (13.28а)$$

Опытные данные и теоретические соображения указывают, что состоянию электрона в валентной зоне в точке $\boldsymbol{k}=0$ соответствует p-образное неприводимое представление Γ_{15}, а состоянию электрона в точке $\boldsymbol{k}=0$ в зоне проводимости — s-образное неприводимое единичное представление Γ_1.

При одновременном рассмотрении зоны проводимости и валентной восемь базисных функций в нулевом приближении целесообразно выбрать следующим образом:

$$iS\mathrm{v}_2,\quad \frac{1}{\sqrt{2}}(x-iy)\,\mathrm{v}_1,\quad z\mathrm{v}_2,\quad \frac{1}{\sqrt{2}}(x+iy)\,\mathrm{v}_1, \quad (13.29)$$

$$iS\mathrm{v}_1,\quad -\frac{1}{\sqrt{2}}(x+iy)\,\mathrm{v}_2,\quad z\mathrm{v}_1,\quad \frac{1}{\sqrt{2}}(x-iy)\,\mathrm{v}_2. \quad (13.29а)$$

Здесь S — сферически симметричная функция представления Γ_1. Функции, стоящие друг под другом в (13.29) и (13.29а), соответствуют, как мы увидим ниже, вырожденным состояниям электрона.

Структура базисных функций (13.29), (13.29а) связана с соображениями, аналогичными тем, которые обсуждались в связи с выбором функций (13.16), (13.16а). Базисные функции (13.29), (13.29а) обеспечивают наиболее простой вид матрицы 8×8 секулярного уравнения (13.7).

Выберем волновой вектор электрона $\boldsymbol{k}$ в направлении оси z ($k_z=k$); тогда матрица 8×8 секулярного уравнения, соответствующая оператору в фигурных скобках уравнения (13.28) на волновых функциях (13.29), (13.29а), имеет вид

$$\begin{pmatrix}\mathscr{H} & 0\\ 0 & \mathscr{H}\end{pmatrix},$$

где

$$\mathscr{H}=\begin{vmatrix}\varepsilon_s & 0 & kP & 0\\ 0 & \varepsilon_p-\Delta/3 & \sqrt{2}\Delta/3 & 0\\ kP & \sqrt{2}\Delta/3 & \varepsilon_p & 0\\ 0 & 0 & 0 & \varepsilon_p+\Delta/3\end{vmatrix}. \quad (13.30)$$

Здесь положительная константа

$$\Delta = \frac{3\hbar i}{4m^2c^2}\left\langle x\left|\frac{\partial V}{\partial x}\hat{p}_y - \frac{\partial V}{\partial y}\hat{p}_x\right|y\right\rangle \qquad (13.30а)$$

равна спин-орбитальному расщеплению в валентной зоне.

Вещественная величина

$$P = -i\left(\frac{\hbar}{m}\right)\langle S|\hat{p}_z|z\rangle \qquad (13.30б)$$

характеризует «взаимодействие» между валентной зоной и зоной проводимости, ε_s и ε_p — энергии, соответствующие краю ($\boldsymbol{k}=0$) зоны проводимости и валентной зоны (без учета спин-орбитального смещения).

Матрица 8×8 (13.30) квазидиагональна. Левый верхний блок ее $\mathscr{H}$, построенный на базисных функциях (13.29), совпадает с правым нижним блоком, построенным на функциях (13.29а).

Для получения матрицы (13.30) были использованы соображения симметрии кристалла (T_d), аналогичные тем, которыми мы пользовались при определении структуры матрицы (13.15). Кроме того, была использована ортонормированность спиновых функций $\nu_1(s)$ и $\nu_2(s)$.

Например, (1, 3)-й член матрицы (13.30) равен

$$\left\langle iS\nu_2\left|\frac{\hat{p}^2}{2m}+V+\frac{\hbar}{m}(k\hat{p}_z)+\frac{\hbar}{4m^2c^2}\left[\left(\frac{\partial V}{\partial y}\hat{p}_z-\frac{\partial V}{\partial z}\hat{p}_y\right)\hat{\sigma}_1+\right.\right.\right.$$
$$\left.\left.\left.+\left(\frac{\partial V}{\partial z}\hat{p}_x-\frac{\partial V}{\partial x}\hat{p}_z\right)\hat{\sigma}_2+\left(\frac{\partial V}{\partial x}\hat{p}_y-\frac{\partial V}{\partial y}\hat{p}_x\right)\hat{\sigma}_3\right]\right|z\nu_2\right\rangle.$$

Матричный элемент от первых двух слагаемых написанного выше оператора равен нулю из соображений симметрии: $\left\langle S\left|\frac{\hat{p}^2}{2m}+V\right|z\right\rangle=0$; матричный элемент первых двух слагаемых в квадратной скобке равен нулю ввиду (10.13) и ортогональности спиновых функций; матричный элемент последнего слагаемого в квадратной скобке равен нулю из соображений симметрии $\left(\left\langle S\left|\frac{\partial V}{\partial x}\hat{p}_y-\frac{\partial V}{\partial y}\hat{p}_x\right|z\right\rangle=0\right)$. Таким образом, остается только матричный элемент третьего слагаемого, так что

$$(1,\,3)\text{-й член} = -ik\left(\frac{\hbar}{m}\right)\langle S|\hat{p}_z|z\rangle = kP,$$

если воспользоваться значением (13.30б).

Матрица (13.30) получена при специальном предположении, что волновой вектор электрона $\boldsymbol{k}$ направлен вдоль оси z. При произвольном направлении $\boldsymbol{k}$ его ориентация задается эйлеровыми углами (см. рис. IV.24): β — полярным углом, α — азимутальным углом, $\gamma=0$.

Координатные базисные функции x, y, z, согласно (10.24) и (10.24а), преобразуются следующим образом:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\alpha\cos\beta & \sin\alpha\cos\beta & -\sin\beta \\ -\sin\alpha & \cos\alpha & 0 \\ \cos\alpha\sin\beta & \sin\alpha\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{13.31}$$

Спиновые функции $v_1(s)$, $v_2(s)$, как следует из (10.30), преобразуются так:

$$\begin{pmatrix} v_1' \\ v_2' \end{pmatrix} = \begin{pmatrix} \cos\frac{\beta}{2} e^{-i\alpha/2} & \sin\frac{\beta}{2} e^{i\alpha/2} \\ -\sin\frac{\beta}{2} e^{-i\alpha/2} & \cos\frac{\beta}{2} e^{i\alpha/2} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{13.31а}$$

В гл. II, § 6 мы отметили, что линейному преобразованию базисных функций соответствует преобразование подобия для соответствующих матриц. С другой стороны, в Приложении 3, п. 4 показано, что матричные уравнения инвариантны относительно преобразования подобия; следовательно, секулярное уравнение при преобразовании подобия имеет те же корни. Мы можем поэтому для секулярного уравнения использовать специальный вид матрицы (13.30).

Таким образом, поправки к энергии ε' во втором приближении теории возмущений можно определить из секулярного уравнения:

$$\begin{vmatrix} \varepsilon_s - \varepsilon' & 0 & kP & 0 \\ 0 & \varepsilon_p - \frac{\Delta}{3} - \varepsilon' & \frac{\sqrt{2}}{3}\Delta & 0 \\ kP & \frac{\sqrt{2}}{3}\Delta & \varepsilon_p - \varepsilon' & 0 \\ 0 & 0 & 0 & \varepsilon_p + \frac{\Delta}{3} - \varepsilon' \end{vmatrix} = 0. \tag{13.32}$$

Поскольку матрица секулярного уравнения (13.30) имеет структуру из двух одинаковых блоков $\mathscr{H}$, корни уравнения (13.32) двукратно вырождены. Разлагая определитель (13.32) по элементам четвертой строки (столбца), получим

$$\left(\varepsilon_p + \frac{\Delta}{3} - \varepsilon'\right)\left[(\varepsilon_s - \varepsilon')\left(\varepsilon_p - \frac{\Delta}{3} - \varepsilon'\right)(\varepsilon_p - \varepsilon') - \right.$$
$$\left. - k^2P^2\left(\varepsilon_p - \frac{\Delta}{3} - \varepsilon'\right) - (\varepsilon_s - \varepsilon')\frac{2\Delta^2}{9}\right] = 0, \tag{13.33}$$

где в квадратных скобках стоит определитель третьего порядка, полученный при вычеркивании 4-й строки и 4-го столбца. Если вести отсчет энергии от верхнего края валентной зоны, т. е. положить

$$\varepsilon_p + \frac{\Delta}{3} = 0, \quad \text{т. е.} \quad \varepsilon_p = -\frac{\Delta}{3} \tag{13.34}$$

и

$$\varepsilon_s = \varepsilon_G = \text{ширине запрещенной зоны},$$

то вместо (13.33) получим

$$\varepsilon' = 0, \tag{13.35}$$

$$\varepsilon'(\varepsilon' - \varepsilon_G)(\varepsilon' + \Delta) - k^2P^2\left(\varepsilon' + \frac{2\Delta}{3}\right) = 0. \tag{13.35а}$$

Используя (13.28а), получим отсюда для малых значений k (с точностью до k^2) следующие четыре решения [1]):

$$\varepsilon_{v1} = \frac{\hbar^2k^2}{2m}, \quad \varepsilon_{v2} = \frac{\hbar^2k^2}{2m} - \frac{2P^2k^2}{3\varepsilon_G}, \tag{13.36}$$

$$\varepsilon_{v3} = -\Delta + \frac{\hbar^2k^2}{2m} - \frac{P^2k^2}{3(\varepsilon_G + \Delta)},$$

$$\varepsilon_c = \varepsilon_G + \frac{\hbar^2k^2}{2m} + \frac{P^2k^2}{3}\left(\frac{2}{\varepsilon_G} + \frac{1}{\varepsilon_G + \Delta}\right).$$

Мы можем отождествить ε_{v1} и ε_{v2} с зонами тяжелых и легких дырок, ε_{v3} с ветвью легких дырок, отщепленной в результате спин-орбитального взаимодействия, и ε_c с ветвью электронов в зоне проводимости.

Следует сразу же отметить, что выражение для ε_{v1} не может быть верным, так как оно не имеет даже правильного знака (энергия дырок отрицательна), что свидетельствует о том, что двухзонное приближение, учитывающее только валентную зону и зону проводимости, недостаточно для интерпретации ветви тяжелых дырок; в этом случае необходимо учитывать вышеи нижележащие зоны.

Так как эффективная масса электрона в зоне проводимости почти в 100 раз меньше массы свободного электрона m, то в выражении ε_c слагаемое, пропорциональное P^2, примерно в 100 раз больше слагаемого $\hbar^2k^2/2m$, т. е. безразмерное отношение $P^2m/\hbar^2\varepsilon_G$ очень велико (~ 100).

Если нам известны величины ε_G и Δ, то из сравнения с опытом можно определить постоянную P^2.

Используя то, что $P^2m/\hbar^2\varepsilon_G \gg 1$, получим из (13.36)

$$\varepsilon_{v2} = -\frac{2P^2k^2}{3\varepsilon_G}, \quad \varepsilon_{v3} = -\Delta - \frac{P^2k^2}{3(\varepsilon_G + \Delta)}, \tag{13.36а}$$

$$\varepsilon_c = \varepsilon_G + \frac{P^2k^2}{3}\left(\frac{2}{\varepsilon_G} + \frac{1}{\varepsilon_G + \Delta}\right).$$

Из этих выражений мы видим, что эффективные массы легких частиц по порядку величины равны $\hbar^2\varepsilon_G/P^2$.

Если $\Delta \gg kP$ и ε_G, то уравнение (13.35а) тоже может быть упрощено; разделив обе части его на Δ и пренебрегая в нуле-

[1]) В нулевом приближении пренебрегаем в (13.35а) слагаемым, пропорциональным k^2P^2, и затем ищем поправки порядка k^2.

вом приближении величинами ε_G/Δ и k^2P^2/Δ, получим в нулевом приближении три корня $\varepsilon'_{01}=\varepsilon'_{02}=0$, $\varepsilon'_{03}=-\Delta$; вводя малую поправку ξ и полагая $\varepsilon'_1=\varepsilon_2=\xi$, $\varepsilon'_3=-\Delta+\xi$, получим из (13.35а)

$$\varepsilon_{v2}=\frac{\hbar^2k^2}{2m}+\frac{\varepsilon_G-(\varepsilon_G^2+8P^2k^2/3)^{1/2}}{2},$$

$$\varepsilon_{v3}=-\Delta+\frac{\hbar^2k^2}{2m}-\frac{P^2k^2}{3\Delta}, \qquad (13.36б)$$

$$\varepsilon_c=\frac{\hbar^2k^2}{2m}+\frac{\varepsilon_G+(\varepsilon_G^2+8P^2k^2/3)^{1/2}}{2}.$$

Из этих выражений следует непараболичность закона дисперсии энергии для электронов ε_c и для легких дырок ε_{v2}.

Для малых значений волнового числа k, разлагая корни в (13.36б) с точностью до k^2, получим (13.36) (для $\Delta\gg\varepsilon_G$). Линейные по k члены в выражении для энергии возникают для валентной зоны InSb для четырехкратно вырожденного спинорного неприводимого представления Γ_8 (табл. IV.9). Эти линейные члены появляются во втором приближении теории возмущений при учете члена $(\boldsymbol{k}\hat{\boldsymbol{p}})$ и спин-орбитального взаимодействия, пропорционального $[\nabla V\times\hat{\boldsymbol{p}}]$ (13.28). Заметим, что такие линейные члены по k не возникают в зоне проводимости InSb и в валентной зоне германия (в последнем случае из-за наличия инверсионной симметрии).

Эти линейные члены вызывают в InSb расщепление вырожденных по спину ветвей тяжелых и легких дырок, пропорциональное k.

При малых значениях $\boldsymbol{k}$ преобладают линейные по k члены, а не члены, пропорциональные k^2 или более высоким степеням k; при больших значениях $\boldsymbol{k}$ преобладают члены, пропорциональные k^2. Это приводит к тому, что максимумы энергии в валентной зоне не совпадают с точкой $\boldsymbol{k}=0$, а сдвинуты относительно нее в направлениях [111]. Опыты показывают, что максимумы энергии всего на 0,015 *эв* выше верхнего края валентной зоны для $\boldsymbol{k}=0$ и расположены весьма близко к этой точке.

Аналогично можно показать, что в зоне проводимости антимонида индия возникают в выражении для энергии члены, пропорциональные k^3; они так же ответственны за снятие спинового вырождения в зоне проводимости.

§ 14. Симметрия, связанная с обращением времени

1. В гл. II, § 9, п. 3 было показано, что если взять уравнение, комплексно-сопряженное временно́му уравнению Шредингера с вещественным гамильтонианом $\mathscr{H}$, то состояние с волновой функцией Ψ^* развивается в направлении времени $-t$, точно так

же, как состояние ψ в направлении времени t. Для стационарного состояния ψ и ψ^* описывают вырожденные состояния, соответствующие одной и той же энергии $\mathscr{E}$.

В этом параграфе мы собираемся в рамках квантовой механики более детально исследовать симметрию, связанную с обращением (инверсией) времени. Вопрос этот довольно сложен, поэтому целый ряд положений мы приведем без доказательств (в этих случаях мы будем писать: «можно доказать»). Необходимые доказательства и дополнения читатель может найти в § 18 книги Г. Л. Бир, Г. Е. Пикус. Симметрия и деформационные эффекты в полупроводниках.— М., 1972. В нашем изложении мы будем близко придерживаться текста этой книги.

2. Можно сказать, что операция инверсии времени $\hat{\mathcal{K}}$ превращает волновую функцию $\psi(\boldsymbol{r}, t)$ в новую функцию

$$\hat{\mathcal{K}}\psi(\boldsymbol{r}, t) = \psi^*(\boldsymbol{r}, -t), \tag{14.1}$$

которая удовлетворяет тому же уравнению Шредингера (если гамильтониан вещественен).

Очевидно, что $\hat{\mathcal{K}}^2\psi = \hat{\mathcal{K}}\hat{\mathcal{K}}\psi = \hat{\mathcal{K}}\psi^* = \psi$, т. е.

$$\hat{\mathcal{K}}^2 = 1. \tag{14.2}$$

Для стационарного состояния

$$(\hat{\mathcal{H}} - \mathscr{E})\psi = 0. \tag{14.3}$$

При вещественном гамильтониане $\hat{\mathcal{H}}$, ψ и $\hat{\mathcal{K}}\psi = \psi^*$ — собственные функции уравнения (14.3), соответствующие одному и тому же собственному значению энергии $\mathscr{E}$; это может привести к дополнительному вырождению состояния с энергией $\mathscr{E}$. Если гамильтониан $\hat{\mathcal{H}}$ инвариантен относительно преобразования $\hat{g}$ группы $\boldsymbol{G}$ ($g \in \boldsymbol{G}$), то ψ и $\hat{g}\psi$ — тоже собственные функции уравнения (14.3), соответствующие одной энергии $\mathscr{E}$. И это тоже может привести к вырождению состояния с энергией $\mathscr{E}$. Однако мы не можем применить теорию, развитую в предыдущих параграфах, к случаю симметрии, связанной с инверсией времени, так как оператор инверсии времени $\hat{\mathcal{K}}$ не является линейным оператором; в самом деле,

$$\hat{\mathcal{K}}(c_1\psi_1 + c_2\psi_2) = c_1^*\hat{\mathcal{K}}\psi_1 + c_2^*\hat{\mathcal{K}}\psi_2, \tag{14.4}$$

в то время как для линейного оператора $\hat{g}$, связанного с преобразованием координат,

$$\hat{g}(c_1\psi_1 + c_2\psi_2) = c_1\hat{g}\psi_1 + c_2\hat{g}\psi_2. \tag{14.4а}$$

Легко показать, что оператор $\hat{\mathcal{K}}$ коммутирует со всеми элементами

$\hat{g}$ группы симметрии $\boldsymbol{G}$; имеем

$$\hat{g}\psi_i = \sum_j D_{ji}\psi_j, \tag{14.5}$$

где $D_{ji}(g)$ — неприводимое представление для элемента $\hat{g}$. Тогда

$$\hat{\mathcal{K}}\hat{g}\psi_i = \hat{\mathcal{K}}\sum_j D_{ji}\psi_j = \sum_j D^*_{ji}\psi^*_j,$$

а

$$\hat{g}\hat{\mathcal{K}}\psi_i = \hat{g}\psi^*_i = \sum_j D^*_{ji}\psi^*_j,$$

откуда

$$\hat{\mathcal{K}}\hat{g} = \hat{g}\hat{\mathcal{K}}. \tag{14.6}$$

Собственные функции ψ и $\hat{\mathcal{K}}\psi$, удовлетворяющие уравнению (14.3) при одном и том же собственном значении энергии $\mathscr{E}$, могут быть линейно независимыми и тогда одному значению $\mathscr{E}$ соответствует два независимых набора ортонормированных собственных функций ψ_i и $\hat{\mathcal{K}}\psi_i$. Если же ψ_i и $\hat{\mathcal{K}}\psi_i$ линейно выражаются друг через друга, то

$$\hat{\mathcal{K}}\psi_i = \sum_j T_{ji}\psi_j, \tag{14.7}$$

где $\boldsymbol{T}$ — унитарная матрица, обеспечивающая ортонормированность функций $\hat{\mathcal{K}}\psi_i$ при ортонормированности функций ψ_i. Можно показать, что в случае (14.7) представления $\boldsymbol{D}$ и $\boldsymbol{D}^*$ (14.5) эквивалентны, т. е.

$$\boldsymbol{D}^* = \boldsymbol{T}^{-1}\boldsymbol{D}\boldsymbol{T}. \tag{14.8}$$

В случае, если функции ψ_i и $\hat{\mathcal{K}}\psi_i$ линейно независимы, т. е. не связаны соотношением (14.7), им могут соответствовать либо эквивалентные, либо неэквивалентные представления $\boldsymbol{D}$ и $\boldsymbol{D}^*$ (таким образом, из (14.7) следует (14.8), но из эквивалентности представлений (14.8), вообще говоря, не следует линейная связь (14.7)).

Таким образом, возможны три случая:

а) ψ и $\hat{\mathcal{K}}\psi$ линейно зависимы; представления $\boldsymbol{D}$ и $\boldsymbol{D}^*$ эквивалентны, т. е. $\chi(g) = \chi^*(g)$;

б) ψ и $\hat{\mathcal{K}}\psi$ линейно независимы; $\boldsymbol{D}$ и $\boldsymbol{D}^*$ неэквивалентны, т. е. $\chi(g) \neq \chi^*(g)$; (14.9)

в) ψ и $\hat{\mathcal{K}}\psi$ линейно независимы; D и D^* эквивалентны, т. е. $\chi(g) = \chi^*(g)$.

Здесь $\chi(g)$ — характер представления $D(g)$, $\chi^*(g)$ — характер D^*. Так как линейно независимые волновые функции ψ и $\hat{\mathcal{K}}\psi$ соответствуют одному и тому же значению энергии $\mathscr{E}$, то в случаях

б) и в) инвариантность по отношению к обращению времени приводит к дополнительному вырождению. Поэтому практически важно различать случаи (14.9).

Докажем, что в случаях а) и в), когда представления $\boldsymbol{D}$ и $\boldsymbol{D}^*$ эквивалентны, т. е. связаны соотношением (14.8), необходимым и достаточным условием вещественности $\boldsymbol{D}$ является требование $\boldsymbol{T}=\tilde{\boldsymbol{T}}$ (из унитарности $\boldsymbol{T}$ следует: $\tilde{\boldsymbol{T}}=\boldsymbol{T}^{*-1}$). Если же $\boldsymbol{T}=-\tilde{\boldsymbol{T}}$, то представление $\boldsymbol{D}$ существенно комплексно, т. е. не может быть каким-либо преобразованием подобия приведено к вещественному виду.

Беря комплексно-сопряженное от (14.8), получим

$$\boldsymbol{D}=\boldsymbol{T}^{*-1}\boldsymbol{D}^*\boldsymbol{T}^*=\boldsymbol{T}^{*-1}\boldsymbol{T}^{-1}\boldsymbol{D}\boldsymbol{T}\boldsymbol{T}^*=(\boldsymbol{T}\boldsymbol{T}^*)^{-1}\boldsymbol{D}(\boldsymbol{T}\boldsymbol{T}^*)$$

или

$$(\boldsymbol{T}\boldsymbol{T}^*)\boldsymbol{D}=\boldsymbol{D}(\boldsymbol{T}\boldsymbol{T}^*),$$

т. е. матрица $\boldsymbol{T}\boldsymbol{T}^*$ коммутирует со всеми матрицами $\boldsymbol{D}(g)$; тогда (по первой лемме Шура [1])) она кратна единичной матрице $\boldsymbol{I}$:

$$\boldsymbol{T}\boldsymbol{T}^*=c\boldsymbol{I},\quad \boldsymbol{T}=c\boldsymbol{T}^{*-1}=c\tilde{\boldsymbol{T}},$$

так как $\boldsymbol{T}$—унитарная матрица. Отсюда $\tilde{\boldsymbol{T}}=c\boldsymbol{T}$ и, следовательно, $\boldsymbol{T}=c\tilde{\boldsymbol{T}}=c^2\boldsymbol{T}$, откуда $c=\pm 1$. Итак, возможны два случая: первый $\boldsymbol{T}=\tilde{\boldsymbol{T}}$ и второй $\boldsymbol{T}=-\tilde{\boldsymbol{T}}$. Если $\boldsymbol{D}(g)$ вещественно, то из (14.8) $\boldsymbol{T}\boldsymbol{D}=\boldsymbol{D}\boldsymbol{T}$, тогда по первой лемме Шура $\boldsymbol{T}=b\boldsymbol{I}$ и $\boldsymbol{T}=\tilde{\boldsymbol{T}}$, т. е. реализуется первый случай. Можно доказать, что условие $\boldsymbol{T}=\tilde{\boldsymbol{T}}$ не только необходимо, но и достаточно для того, чтобы $\boldsymbol{D}(g)$ было вещественно. Таким образом, мы доказали высказанные выше утверждения. Пользуясь этим результатом, докажем, что в случае а) (14.9), когда имеется линейная связь между ψ и $\hat{\mathcal{K}}\psi$ (14.7) и, следовательно $\boldsymbol{T}$ в (14.8) совпадает с $\boldsymbol{T}$ в (14.7), представление $\boldsymbol{D}$ вещественно.

Из (14.7) следует, что

$$\hat{\mathcal{K}}^2\psi_i=\sum_j T^*_{ji}\hat{\mathcal{K}}\psi_j=\sum_{jl}T^*_{ji}T_{lj}\psi_l=\sum_l(TT^*)_{li}\psi_l.$$

Так как $\hat{\mathcal{K}}^2=1$ (14.2), то $(\boldsymbol{T}\boldsymbol{T}^*)_{li}=\delta_{li}$ и, следовательно, $\boldsymbol{T}\boldsymbol{T}^*=\boldsymbol{I}$ или $\boldsymbol{T}=\tilde{\boldsymbol{T}}$ (здесь опять использована унитарность $\boldsymbol{T}$). Таким образом, случаю а) всегда соответствует вещественное представление $\boldsymbol{D}$. И, наоборот, если представление $\boldsymbol{D}$ вещественно, т. е. реализуется случай а), то симметрия по отношению к инверсии времени не приводит к дополнительному вырождению.

Если же представление $\boldsymbol{D}$ (14.5), по которому преобразуются функции, комплексно, то инверсия времени приводит к дополнительному двойному вырождению, независимо от того, яв-

[1]) Бир Г. Л., Пикус Г. Е., § 8.

ляются ли представления $\boldsymbol{D}$ и $\boldsymbol{D}^*$ эквивалентными (случай в)) или неэквивалентными (случай б)).

3. Мы не учитывали пока спина электрона. Член в гамильтониане $\hat{\mathcal{H}}$, учитывающий спин-орбитальное взаимодействие, имеет вид (10.6)

$$\hat{\mathcal{H}}_{s0} = -\frac{i\hbar^2}{4m^2c^2}\hat{\boldsymbol{\sigma}}[\nabla V \times \nabla], \tag{14.10}$$

где $\hat{\sigma} = \{\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3\}$ — спиновые матрицы Паули (10.2). При инверсии времени, т. е. переходе к комплексному сопряжению, (14.10) приобретает вид

$$\hat{\mathcal{H}}^*_{s0} = \frac{i\hbar^2}{4m^2c^2}\hat{\boldsymbol{\sigma}}^*[\nabla V \times \nabla] = -\frac{i\hbar^2}{4m^2c^2}(-\hat{\boldsymbol{\sigma}}^*[\nabla V \times V]). \tag{14.10а}$$

Здесь, как это следует из (10.2),

$$\hat{\sigma}_1^* = \hat{\sigma}_1, \quad \hat{\sigma}_2^* = -\hat{\sigma}_2, \quad \hat{\sigma}_3^* = \hat{\sigma}_3. \tag{14.10б}$$

Для полного гамильтониана, рассматриваемого как функционал от $\hat{\sigma}_i$

$$\hat{\mathcal{H}}^*(\hat{\sigma}_i) = -\hat{\mathcal{H}}(-\hat{\sigma}_i^*). \tag{14.11}$$

Уравнение Шредингера—Паули для стационарных состояний имеет вид

$$[\hat{\mathcal{H}}(\hat{\sigma}_i) - \mathcal{E}]\Psi(\boldsymbol{r}, s) = 0, \tag{14.12}$$

где волновая функция (см. (10.11) и (10.11а))

$$\Psi(\boldsymbol{r}, s) = \sum_{i=1,2} \psi_i(\boldsymbol{r})\, v_i(s), \tag{14.13}$$

или

$$\Psi(\boldsymbol{r}, s) = \begin{pmatrix} \psi_1(\boldsymbol{r}) \\ \psi_2(\boldsymbol{r}) \end{pmatrix}. \tag{14.13а}$$

Операция обращения времени (комплексного сопряжения) (14.12) дает

$$[\hat{\mathcal{H}}^*(\hat{\sigma}_i) - \mathcal{E}]\Psi^*(\boldsymbol{r}, s) = 0. \tag{14.14}$$

Мы не можем теперь (как это было, когда не учитывался спин) сказать, что $\Psi(\boldsymbol{r}, s)$ и $\Psi^*(\boldsymbol{r}, s)$ являются решениями одного и того же уравнения (14.12).

Подвергнем уравнение (14.14) каноническому преобразованию, потребовав, чтобы

$$\boldsymbol{S}\hat{\mathcal{H}}^*(\hat{\sigma}_i)\boldsymbol{S}^{-1} = \boldsymbol{S}\hat{\mathcal{H}}(-\hat{\sigma}_i^*)\boldsymbol{S}^{-1} = \hat{\mathcal{H}}(\hat{\sigma}_i), \tag{14.15}$$

где $\boldsymbol{S}$ — матрица унитарного преобразования, действующая на спиновые матрицы $\hat{\sigma}_i$.

Очевидно, для этого

$$\boldsymbol{S}\hat{\sigma}_1\boldsymbol{S}^{-1} = -\hat{\sigma}_1, \quad \boldsymbol{S}\hat{\sigma}_2\boldsymbol{S}^{-1} = \hat{\sigma}_2, \quad \boldsymbol{S}\hat{\sigma}_3\boldsymbol{S}^{-1} = -\hat{\sigma}_3, \tag{14.16}$$

как это следует из (14.10б).

Воспользовавшись (10.2а), легко проверить, что эти уравнения удовлетворяются, если положить

$$S = S^{-1} = \hat{\sigma}_2. \tag{14.17}$$

Каноническому преобразованию (14.15) соответствует преобразованная волновая функция

$$S\Psi^*(\boldsymbol{r}, s) = \hat{\sigma}_2 \hat{\mathcal{K}}_0 \Psi(\boldsymbol{r}, s), \tag{14.18}$$

где $\hat{\mathcal{K}}_0$ — оператор комплексного сопряжения.

Таким образом, функция $\hat{\sigma}_2 \hat{\mathcal{K}}_0 \Psi(\boldsymbol{r}, s)$ удовлетворяет тому же уравнению (14.12), что и функция $\Psi(\boldsymbol{r}, s)$. Инверсии времени в случае уравнения Шредингера—Паули (14.12) соответствует оператор

$$\hat{\mathcal{K}} = \hat{\sigma}_2 \hat{\mathcal{K}}_0. \tag{14.19}$$

Применяя его дважды, получим

$$\hat{\mathcal{K}}^2 \Psi = \hat{\sigma}_2 \hat{\mathcal{K}}_0 \hat{\sigma}_2 \hat{\mathcal{K}}_0 \Psi = \hat{\sigma}_2 \hat{\mathcal{K}}_0 \sigma_2 \Psi^* = \hat{\sigma}_2 \hat{\sigma}_2^* \Psi = -\Psi,$$

где мы использовали (10.2а); таким образом,

$$\hat{\mathcal{K}}^2 = -1 \tag{14.20}$$

в противоположность (14.2) для бесспинового случая.

Используя (14.20), можно совершенно аналогично тому, как это сделано выше, показать, что случаю а) в (14.9) соответствует необходимое и достаточное условие $\boldsymbol{T} = -\tilde{\boldsymbol{T}}$. Таким образом, при учете спина случаю а) в (14.9) соответствуют эквивалентные, существенно комплексные представления $\boldsymbol{D}(g)$.

И в этом случае, как и в отсутствие спина, инвариантность к инверсии времени не приводит к добавочному вырождению. В тех же случаях, когда спинорные волновые функции преобразуются по двузначному вещественному представлению (случай в) в (14.9)) или по комплексному с комплексными характерами (случай б) в (14.9)), представления удваиваются, т. е. имеет место дополнительное двойное вырождение.

4. Фробениус и Шур показали, что, используя свойства матрицы $\boldsymbol{T}$, можно установить, является ли представление вещественным или комплексным, зная лишь его характеры: если сумма характеров квадратов элементов группы равна числу элементов группы h, то $\boldsymbol{T} = \tilde{\boldsymbol{T}}$ и представление вещественно; если эта сумма равна $-h$, то $\boldsymbol{T} = -\tilde{\boldsymbol{T}}$ и представления комплексны и эквивалентны; наконец, если она равна нулю, то представления комплексны и неэквивалентны.

Эти результаты отражены в следующей таблице, учитывающей влияние инверсии времени на вырождение.

Таблица IV.17 [1])

Критерий Фробениуса—Шура	Вырождение	Соотношения между D и D^*	
		В отсутствие спина (D — обычные представления)	При наличии спина (D — спинорные представления)
а) $\sum_g \chi(g^2) = \hat{\mathcal{K}}^2 h$	Нет дополнительного вырождения	D и D^* могут быть сделаны вещественными и одинаковыми	D и D^* эквивалентны, но существенно комплексны (т. е. не могут быть сделаны вещественными)
б) $\sum_g \chi(g^2) = 0$	Имеется дополнительное двойное вырождение	D и D^* не эквивалентны $\chi(g) \neq \chi^*(g)$	
в) $\sum_g \chi(g^2) = -\hat{\mathcal{K}}^2 h$		D и D^* эквивалентны, но существенно комплексны (т. е. не могут быть сделаны вещественными)	D и D^* могут быть сделаны вещественными и одинаковыми

[1]) Величина $\hat{\mathcal{K}}^2 = 1$ в отсутствие спина (14.2) и $\hat{\mathcal{K}}^2 = -1$ при наличии спина (14.20).

Продемонстрируем применение этой таблицы на простом примере точечной группы C_3 (табл. II.8). Критерий Фробениуса — Шура для неприводимого представления Γ_2 дает

$$\sum_g \chi(g^2) = \chi(E^2) + \chi(C_3^2) + x(C_3^4) = \chi(E) + \chi(C_3^2) + \chi(C_3) =$$
$$= 1 + \omega^2 + \omega = 1 + e^{-2\pi i/3} + e^{2\pi i/3} = 1 + 2\cos\frac{2\pi}{3} = 0.$$

Из приведенной выше таблицы видно, что для Γ_2 (и аналогично для Γ_3) имеется дополнительное двойное вырождение, связанное с симметрией по отношению к инверсии времени. Это означает, что состояния Γ_2 и Γ_3 объединяются в одно неприводимое двукратно вырожденное состояние (в табл. II.8 они поэтому объединены).

5. Непосредственное применение критерия Фробениуса—Шура (табл. IV.17) к исследованию влияния инверсии времени на уровни энергии электрона в кристалле невозможно, так как суммирование в $\sum_g \chi(g^2)$ должно распространяться на все элементы

группы симметрии системы, а число элементов пространственной группы кристалла (практически) бесконечно.

Херрингу удалось преобразовать критерий Фробениуса — Шура применительно к энергетическим зонам в кристалле, выразив его через характеры элементов, относящихся к группе волнового вектора[1]). Критерий Херринга, позволяющий различать случаи а), б) и в) табл. IV.17, имеет вид

$$\sum_{g_0}\chi(g_0^2)=\begin{cases}\hat{\mathcal{K}}^{e^2}n, & \text{случай а),}\\ 0, & \text{случай б),}\\ -\hat{\mathcal{K}}^{e^2}n, & \text{случай в).}\end{cases} \tag{14.21}$$

Здесь g_0 — элемент пространственной группы кристалла $\boldsymbol{G}$, не содержащий тривиальных трансляций, преобразующий волновой вектор $\boldsymbol{k}$ в $-\boldsymbol{k}$; поэтому g_0^2 преобразует $\boldsymbol{k}\to\boldsymbol{k}$, т. е. в один из элементов группы волнового вектора $\boldsymbol{G}_k$, который для несимморфных групп, вообще говоря, уже может содержать тривиальную трансляцию, n — число таких элементов. Заметим, что $\chi(g_0^2)$ — характер исследуемого представления группы волнового вектора $\boldsymbol{G}_k$ для элемента g_0^2. В конце § 8 мы исследовали энергетический спектр In Sb. Мы показали, что если центру бриллюэновской зоны соответствует трехкратно вырожденное состояние Γ_{15}, то на оси Δ оно расщепляется следующим образом:

$$\Gamma_{15}=\Delta_1+\Delta_3+\Delta_4,$$

где Δ_1, Δ_3, Δ_4 — неприводимые одномерные представления группы волнового вектора $\boldsymbol{k}_\Delta$. Далее мы отметили, что состояния Δ_3 и Δ_4 из-за дополнительной симметрии, связанной с инверсией времени, не расщепляются, образуя двукратно вырожденное состояние.

Исследуем этот случай посредством критерия Херринга (14.21). Элементы группы $\boldsymbol{T}_d$, точечной подгруппы пространственной группы кристалла In Sb, даны в табл. IV.2. Легко видеть, что из них четыре элемента преобразуют $\boldsymbol{k}_\Delta\to-\boldsymbol{k}_\Delta$:

$$R_3(\bar{x}y\bar{z})=C_4^2,\quad R_4(\bar{x}\bar{y}z)=C_4'^2,$$
$$R_{13}(\bar{x}z\bar{y})=JC_4,\quad R_{14}(\bar{x}\bar{z}y)=JC_4'.$$

Далее

$$R_3^2=R_1(xyz)=E,\quad R_4^2=R_1(xyz)=E,$$
$$R_{13}^2=R_2(x\bar{y}\bar{z})=C_4^2,\quad R_{14}^2=R_2(x\bar{y}\bar{z})=C_4^2.$$

---

[1]) См. статью: Херринг К. Влияние симметрии относительно инверсии времени на энергетические зоны кристаллов./В Сб. Нокс. Р., Голд А. Симметрия в твердом теле.— М., 1970, с. 243).

Обращаясь к табл. IV.4 характеров группы волнового вектора $\boldsymbol{k}_\Delta$, имеем для представлений Δ_3 и Δ_4

$$\sum_{g_0}\chi(g_0^2)=\chi(R_3^2)+\chi(R_4^2)+\chi(R_{13}^2)+\chi(R_{14}^2)=\chi(E)+\chi(E)+$$
$$+\chi(C_4^2)+\chi(C_4^2)=1+1-1-1=0,$$

т. е., согласно (14.21), реализуется случай б). Из (14.21) следует, что как при учете спина, так и в бесспиновом случае состояния Δ_3 и Δ_4 двукратно вырождены, т. е. вдоль оси Δ не расщепляются.

Если пространственная группа кристалла содержит инверсию J (Ge, Si), то для наиболее общего положения волнового вектора $\boldsymbol{k}$ единственный элемент $g_0=J$, тогда $\sum_{g_0}\chi(g_0^2)=\chi(J^2)=$ $=\chi(E)=1$ и в (14.21) реализуется случай а), что при учете спина приводит к двукратному вырождению; это обстоятельство было уже отмечено при обсуждении выражений (13.23) и (13.26).

§ 15. Структура энергетических зон некоторых полупроводников

В настоящее время на основании многочисленных опытов и теоретических исследований определена энергетическая структура ряда полупроводников: германия, кремния, антимонида индия и др. Выяснилось, что практически ни в одном случае энергетический спектр электронов и дырок проводимости не имеет простого параболического характера $\varepsilon=p^2/2m^*$, где m^* — скалярная эффективная масса.

Наибольшую информацию об энергетической структуре полупроводников дали опыты по циклотронному резонансу, поглощению света и сопротивлению в магнитном поле. На некоторых из этих опытов мы остановимся в следующих главах.

Одним из наиболее эффективных методов расчета энергетических зон в полупроводниках является метод *ортогонализованных плоских волн* (К. Херринг, 1940), который удачно объединяет приближения почти свободных (см. § 5) и сильно связанных электронов (см. § 7)[1]. Серьезную помощь в классификации состояний и выборе волновых функций электронов (дырок) проводимости оказывают соображения теории групп, т. е. свойства симметрии, рассмотренные в предыдущих параграфах. Для расчета наряду с аналитическими методами применялись ЭВМ.

---

[1]) Общее представление о различных методах расчета зонной структуры в твердых телах с соответствующей библиографией можно получить из книги: Цидильковский И. М. Электроны и дырки в полупроводниках.—М., 1972, гл. II.

Такие расчеты для германия и кремния были проведены Херманом с сотрудниками (1953). При расчете пришлось ограничиться вычислением энергии в некоторых симметричных точках зоны Бриллюэна, для которых выкладки сильно упрощаются. Значение энергии для промежуточных точек получено интерполированием.

Для германия и кремния бриллюэновская зона имеет форму четырнадцатигранника, изображенного на рис. IV.22. Оси x, y, z проходят через центры шести квадратов и совпадают с кристаллографическими осями [100], [010], [001] и т. д. Оси, направленные к точкам L — центрам шестиугольников, совпадают с осями $[11\bar{1}]$, $[111]$ и т. д. (точка L на рис. IV.22 имеет координаты: $\frac{2\pi}{a}\left[\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}\right]$, т. е. $k_x=k_y=k_z=\frac{\pi}{a}$, где a — ребро куба прямой решетки).

На рис. IV.28 дана качественная картина зонной структуры германия (*а*) и кремния (*б*) (без учета спин-орбитального расщепления); оси $\Gamma\Delta X$ и $\Gamma\Lambda L$ соответствуют направлениям [100] и [111]; заштрихованная часть — запрещенная зона.

Как показывают расчеты и опыт для Ge и Si для дырок в центре валентной зоны ($\boldsymbol{k}=0$) реализуется трехкратно вырожденное (без учета спина) состояние Γ'_{25}.

Центру зоны проводимости ($\boldsymbol{k}=0$) для Ge соответствует невырожденное состояние с наименьшей энергией Γ'_2, а для Si — трехкратно вырожденное состояние Γ_{15}.

Пользуясь табл. III.5 совместности, легко определить характер расщепления состояний электрона (дырки) при переходе из центра Г на оси Δ или Λ (при выходе на поверхность зоны Бриллюэна нужна, ввиду несимморфности пространственной группы Ge и Si, некоторая осторожность; см. конец § 9).

Ширина запрещенной зоны у германия $\varepsilon_G=\mathscr{E}(L_1)-\mathscr{E}(\Gamma'_{25})=0{,}8$ *эв*; у кремния — $\varepsilon_G=\mathscr{E}(\Delta_1)-\mathscr{E}(\Gamma'_{25})=1{,}1$ *эв*.

Зона проводимости кремния имеет шесть симметрично расположенных минимумов энергии в точках на осях Δ, т. е. в направлениях [100]. В зоне проводимости германия имеется восемь минимумов энергии на границе зоны Бриллюэна в точках L, т. е. в направлениях [111][1]). Изоэнергетические поверхности энергии $\varepsilon(\boldsymbol{k})=\text{const}$ вблизи этих минимумов имеют вид эллипсоидов вращения (§ 3, п. 3) с осью симметрии, направленной у кремния вдоль [100], а у германия — вдоль [111].

Зависимость энергии вблизи минимума от квазиимпульса $\boldsymbol{p}=\hbar\boldsymbol{k}$ в главных осях имеет вид

$$\varepsilon(\boldsymbol{p})=\frac{p_1^2}{2m_1}+\frac{p_2^2}{2m_2}+\frac{p_3^2}{2m_3}=\frac{p_1^2+p_2^2}{2m_\perp}+\frac{p_3^2}{2m_\parallel}. \qquad (15{,}1)$$

[1]) Так как каждый минимум принадлежит двум зонам Бриллюэна (лежит на их границе), то на одну зону приходится четыре минимума.

Здесь $m_1 = m_2 = m_\perp$ — поперечная, а $m_3 = m_\parallel$ — продольная эффективные массы.

Из опытов по циклотронному резонансу и магнетопоглощению света для германия

$$m_\perp = 0{,}082m, \quad m_\parallel = 1{,}59m, \quad m_\parallel/m_\perp = 19, \tag{15.2}$$

где $m = 9{,}11 \cdot 10^{-28}$ *г* — масса свободного электрона.

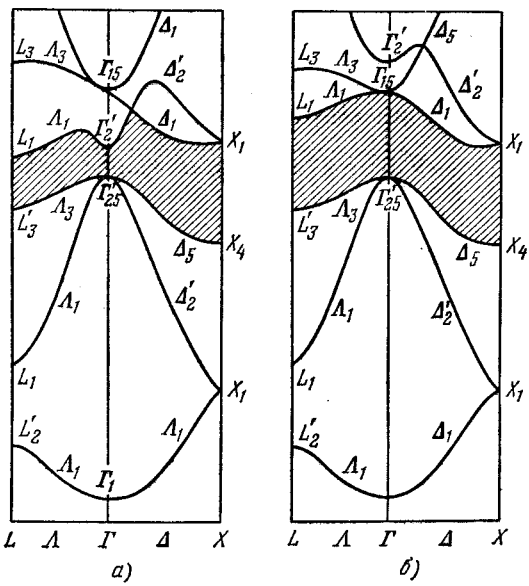


Рис. IV. 28.

Для кремния

$$m_\perp = 0{,}19m, \quad m_\parallel = 0{,}92m, \quad m_\parallel/m_\perp = 4{,}8. \tag{15.3}$$

Мы уже упоминали, что в Ge и Si в центре валентной зоны ($\boldsymbol{k} = 0$) реализуется неприводимое представление Γ_{25}', к которому приложимы соображения о виде энергетического спектра, развитые в § 13, п. 3.

Энергетический спектр дырок, как это следует из (13.23) и (13.26), состоит из трех ветвей (зон):

$$\varepsilon_{v1} = -\frac{\hbar^2}{2m}\left[Ak^2 - \sqrt{B^2k^4 + C^2\left(k_x^2k_y^2 + k_x^2k_z^2 + k_y^2k_z^2\right)}\right], \tag{15.4}$$

$$\varepsilon_{v2} = -\frac{\hbar^2}{2m}\left[Ak^2 + \sqrt{B^2k^4 + C^2\left(k_x^2k_y^2 + k_x^2k_z^2 + k_y^2k_z^2\right)}\right], \tag{15.4а}$$

$$\varepsilon_{v3} = -\Delta - \frac{\hbar^2}{2m}Ak^2, \tag{15.4б}$$

где мы вместо постоянных A, B и C, входящих в (13.23), ввели новые безразмерные постоянные A, B и C, равные старым, деленным на $(-\hbar^2/2m)$. Здесь ε_{v1}—зона тяжелых дырок, ε_{v2}—зона легких дырок и ε_{v3}—зона дырок, отщепленная от верхнего края валентной зоны на величину Δ из-за спин-орбитального взаимодействия. Только для последней отщепленной зоны имеет место простой параболический закон $\varepsilon_{v3}+\Delta \sim k^2$ и поверхности постоянной энергии—сферы. В случае (15.4) и (15.4а) поверхности постоянной энергии имеют вид *гофрированных* поверхностей, сечение которых плоскостью (100) изображено на рис. IV.27 [1]).

Вводя в $\boldsymbol{k}$-пространстве вместо прямоугольных сферические координаты [2]), получим вместо (15.4)

$$\varepsilon_{v1} = -\frac{\hbar^2 k^2}{2m}\left[A - \sqrt{B^2 + C^2 \sin^2\vartheta\,(\sin^2\vartheta \sin^2\varphi \cos^2\varphi + \cos^2\vartheta)}\right]. \tag{15.5}$$

При изотропном усреднении по углам получим приближенно

$$\varepsilon_{v1} \approx -\frac{\hbar^2 k^2}{2m}\left[A - \sqrt{B^2 + \frac{C^2}{5}}\right] \tag{15.5а}$$

и аналогичное выражение со знаком «плюс» перед корнем для ε_{v2}. Выражение (15.5а) и аналогичное для ε_{v2} позволяют ввести скалярные эффективные массы

$$m_{v1} = \frac{m}{A - \sqrt{B^2 + C^2/5}} \text{ и } m_{v2} = \frac{m}{A + \sqrt{B^2 + C^2/5}} \tag{15.5б}$$

для тяжелых и легких дырок.

Исследование циклотронного резонанса в германии дало следующие значения постоянных. $A = 13{,}1 \pm 0{,}4$, $B = 8{,}3 \pm 0{,}6$, $C = 12{,}5 \pm 0{,}5$ и $\Delta = 0{,}3$ *эв*; эти же константы для кремния равны: $A = 4{,}0 \pm 0{,}1$, $B = 1{,}1 \pm 0{,}4$, $C = 4{,}1 \pm 0{,}4$ и $\Delta = 0{,}04$ *эв*.

Используя эти численные значения, получим из (15.5а) для эффективных масс тяжелых и легких дырок в

$$\text{Ge:}\quad m_{v1} = 0{,}33m,\quad m_{v2} = 0{,}04m,\quad m_{v1}/m_{v2} = 8{,}0, \tag{15.6}$$

и в

$$\text{Si:}\quad m_{v1} = 0{,}56m,\quad m_{v2} = 0{,}16m,\quad m_{v1}/m_{v2} = 3{,}5.$$

В § 13, п. 4 мы исследовали структуру энергетических зон антимонида индия—полупроводника с узкой запрещенной зоной. В этом случае энергетические зоны обладают двумя характерными особенностями: очень малыми эффективными массами носителей тока и отступлением закона дисперсии энергии от простой параболичности. Там же мы привели численные значения основных параметров, характеризующих зонную структуру InSb.

[1]) Значения констант A, B, C взяты для Ge.

[2]) Полярную ось направим по оси симметрии эллипсоида, т. е. по k_z.

Аналогичными свойствами обладают и некоторые другие соединения элементов III и V групп периодической системы элементов: арсенид индия (InAs), фосфид индия (InP), арсенид галлия (GaAs), сурьмянистый алюминий (AlSb), фосфид алюминия (AlP). Соединения этого типа обозначаются как $A^{III}B^{IV}$.

В то же время, некоторые другие полупроводники — соединения типа $A^{III}B^{V}$, например фосфид галия (GaP), имеют зоны, для которых экстремумы энергии электронов проводимости и дырок расположены в разных $\boldsymbol{k}$-точках зоны Бриллюэна [1]).

---

[1]) Мы не будем приводить численные значения параметров, характеризующих зонную структуру соединений $A^{III}B^{V}$ и обсуждать их важные технические применения. Для детального ознакомления с этими вопросами читатель может обратиться к монографии: Маделунг О. Физика полупроводниковых соединений элементов III и V групп.—М.: Мир, 1967.

ГЛАВА V

ЛОКАЛИЗОВАННЫЕ СОСТОЯНИЯ ЭЛЕКТРОНА В КРИСТАЛЛЕ

§ 1. Функции Ванье. Движение электрона в поле примеси

1. При изучении поведения электронов проводимости в полупроводниках нередко возникает ситуация, когда на электрон, помимо периодического потенциала идеального кристалла, действует некоторое дополнительное поле, которое может приводить к локализованным (связанным) состояниям электрона в решетке. Такое положение имеет место, например, при движении электрона проводимости в поле постороннего атома или иона, внедренного в идеальную решетку кристалла (*примесный атом* или *ион*).

Наличие свободной поверхности у кристалла также эквивалентно существованию некоторого дополнительного поля вблизи поверхности. Таким же дополнительным полем является кулоновское поле дырки, действующее на электрон, которое в некоторых случаях приводит к связанным состояниям электрона и дырки, называемых *экситоном*. Наконец, дополнительное поле, создаваемое поляризацией ионного кристалла, которая вызвана действием самого электрона на решетку, приводит к так называемым *поляронным состоянием*.

В некоторых случаях для описания локализованных состояний электрона в решетке целесообразно использовать в качестве нулевого приближения не функции Блоха

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = u_{n\boldsymbol{k}}(\boldsymbol{r})\, e^{i\boldsymbol{k}\boldsymbol{r}}, \tag{1.1}$$

где n — номер разрешенной зоны энергии, $\boldsymbol{k}$ — волновой вектор, а так называемые *функции Ванье* (1.3), определенные ниже. Функции Блоха (1.1) периодичны в $\boldsymbol{k}$-пространстве с периодами обратной решетки $\boldsymbol{b}_i$ ($i = 1, 2, 3$); это означает, что $\psi_{n\boldsymbol{k}}(\boldsymbol{r})$ может быть разложено в $\boldsymbol{k}$-пространстве в ряд Фурье:

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{N}} \sum_{l} \varphi_n(\boldsymbol{a}_l, \boldsymbol{r})\, e^{i\boldsymbol{k}\boldsymbol{a}_l}. \tag{1.2}$$

В самом деле, сумма в правой части не меняется при замене $\boldsymbol{k}$ на $\boldsymbol{k}+\boldsymbol{b}_i$: $\exp[i(\boldsymbol{k}+\boldsymbol{b}_i)\boldsymbol{a}_l]=\exp[i\boldsymbol{k}\boldsymbol{a}_l]\exp[i\boldsymbol{b}_i\boldsymbol{a}_l]=\exp[i\boldsymbol{k}\boldsymbol{a}_l]\times$ $\times\exp[i(2\pi\times\text{целое число})]=\exp[i\boldsymbol{k}\boldsymbol{a}_l]$. Коэффициенты разложения $\varphi_n(\boldsymbol{a}_l, \boldsymbol{r})$, где $\boldsymbol{a}_l$—вектор решетки $\boldsymbol{l}$-го узла, называются функциями Ванье; они осуществляют «узельное представление» электрона в идеальной решетке.

Умножая обе части равенства (1.2) на $\exp[-i\boldsymbol{k}\boldsymbol{a}_m]$ и суммируя по $\boldsymbol{k}$, получим, используя (П.6.8),

$$\varphi_n(\boldsymbol{a}_m, \boldsymbol{r})=\frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}}e^{-i\boldsymbol{k}\boldsymbol{a}_m}\psi_{n\boldsymbol{k}}(\boldsymbol{r})=\frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}}u_{n\boldsymbol{k}}e^{i\boldsymbol{k}(\boldsymbol{r}-\boldsymbol{a}_m)}. \quad (1.3)$$

Так как $u_{n\boldsymbol{k}}(\boldsymbol{r})=u_{n\boldsymbol{k}}(\boldsymbol{r}-\boldsymbol{a}_m)$, то $\varphi_n(\boldsymbol{a}_m, \boldsymbol{r})=\varphi_n(\boldsymbol{r}-\boldsymbol{a}_m)$.

Покажем, что каждая из N функций Ванье $\varphi_n(\boldsymbol{r}-\boldsymbol{a}_m)$ ($\boldsymbol{m}=1, 2, \ldots, N$) локализована вблизи своего узла $\boldsymbol{m}$. Для упрощения доказательства рассмотрим простую кубическую решетку и аппроксимируем функцию Блоха плоской волной $\psi_{n\boldsymbol{k}}(\boldsymbol{r})=\frac{1}{\sqrt{V}}e^{i\boldsymbol{k}\boldsymbol{r}}$; в этом случае (1.3) равно

$$\varphi_n(\boldsymbol{r}-\boldsymbol{a}_m)=C\sum_{k_x}e^{ik_x\xi}\sum_{k_y}e^{ik_y\eta}\sum_{k_z}e^{ik_z\zeta}.$$

Здесь C—нормировочная константа волновой функции φ_n; ξ, η, ζ—прямоугольные составляющие радиуса-вектора $\boldsymbol{\rho}=\boldsymbol{r}-\boldsymbol{a}_m$ и $k_i=2\pi g_i/Ga$ ($-G/2\leqslant g_i\leqslant G/2$) [1]), как это следует из (IV.3.3) для простой кубической решетки с ребром куба, равным a.

Суммируя, как геометрическую прогрессию, сумму по k_x (по g), получим

$$\sum_{k_x}e^{ik_x\xi}=1+\sum_{g=1}^{G/2}e^{i\frac{2\pi\xi}{Ga}g}+\sum_{g=-1}^{-G/2}e^{i\frac{2\pi\xi}{Ga}g}=\frac{e^{i\frac{\pi\xi}{a}}e^{i\frac{2\pi\xi}{Ga}}-e^{-i\frac{\pi\xi}{a}}}{e^{i\frac{2\pi\xi}{Ga}}-1}. \quad (1.3a)$$

При достаточно больших размерах основной области кристалла можно считать, что $\xi/Ga\ll 1$, поэтому $\exp i\frac{2\pi}{Ga}\xi\approx 1+i\frac{2\pi}{Ga}\xi$; тогда (1.3а) (с точностью до константы) равно

$$\sum_{k_x}e^{ik_x\xi}\approx\frac{\sin\frac{\pi\xi}{a}}{\frac{\pi\xi}{a}}.$$

Это выражение имеет максимальное значение, равное единице при $\xi=0$; при увеличении ξ оно быстро убывает, осциллируя. Таким образом, функция Ванье $\varphi_n(\boldsymbol{r}-\boldsymbol{a}_m)$ имеет максимум в точке $\boldsymbol{r}=\boldsymbol{a}_m$, быстро убывая при возрастании $\boldsymbol{r}-\boldsymbol{a}_m$.

[1]) В данном случае для G выбрано большое четное число.

Аналогичное доказательство может быть приведено и в более общем случае—для кристалла произвольной структуры и для общего вида блоховской функции (следует только считать, что $u_{\boldsymbol{k}}(\boldsymbol{r})$ слабо зависит от $\boldsymbol{k}$). В этом смысле выражение (1.2) подобно приближению сильной связи (IV.7.2а), но является точным. Выражение (1.2) имеет существенное преимущество перед (IV.7.2а), так как функции Ванье (1.3) взаимно ортогональны как по номеру зоны n, так и по номеру узла $\boldsymbol{m}$; в самом деле,

$$\int_V \varphi_n^*(\boldsymbol{r}-\boldsymbol{a}_m)\,\varphi_{n'}(\boldsymbol{r}-\boldsymbol{a}_{m'})\,d^3r =$$

$$= \frac{1}{N}\sum_{\boldsymbol{k}\boldsymbol{k}'} e^{i(\boldsymbol{k}\boldsymbol{a}_m-\boldsymbol{k}'\boldsymbol{a}_{m'})}\int \psi_{n\boldsymbol{k}}^*(\boldsymbol{r})\,\psi_{n'\boldsymbol{k}'}(\boldsymbol{r})\,d^3r =$$

$$= \frac{1}{N}\sum_{\boldsymbol{k}\boldsymbol{k}'} e^{i\boldsymbol{k}(\boldsymbol{a}_m-\boldsymbol{a}_{m'})}e^{i(\boldsymbol{k}-\boldsymbol{k}')\boldsymbol{a}_{m'}}\delta_{nn'}\delta_{\boldsymbol{k}\boldsymbol{k}'} =$$

$$= \frac{\delta_{nn'}}{N}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}(\boldsymbol{a}_m-\boldsymbol{a}_{m'})} = \delta_{nn'}\delta_{mm'}, \qquad (1.4)$$

где мы использовали (IV.3.4) и (П.6.8).

2. Пусть наряду с периодическим потенциалом $V(\boldsymbol{r})$ на электрон проводимости действует поле $\mathcal{U}(\boldsymbol{r})$. Уравнение Шредингера в этом случае имеет вид

$$-\frac{\hbar^2}{2m}\nabla^2\Phi_i(\boldsymbol{r}) + [V(\boldsymbol{r})+\mathcal{U}(\boldsymbol{r})]\,\Phi_i(\boldsymbol{r}) = \mathcal{E}_i\Phi_i(\boldsymbol{r}), \qquad (1.5)$$

где $\Phi_i(\boldsymbol{r})$ и $\mathcal{E}_i$—i-я собственная функция и соответствующая ей собственная энергия.

Разложим собственную функцию $\Phi_i(\boldsymbol{r})$ по полной системе ортонормированных функций Ванье (1.3):

$$\Phi_i(\boldsymbol{r}) = \sum_{n'l'} f_{n'}^i(\boldsymbol{a}_{l'})\,\varphi_{n'}(\boldsymbol{r}-\boldsymbol{a}_{l'}). \qquad (1.6)$$

Здесь n'—номер зоны энергии, а $\boldsymbol{a}_{l'}$—вектор решетки узла l', f_n^i—коэффициенты разложения, которые надо определить из (1.5). Подставляя (1.6) в (1.5), умножая слева на $\varphi_n^*(\boldsymbol{r}-\boldsymbol{a}_l)$ и интегрируя по $d^3r = dx\,dy\,dz$, получим

$$\sum_{n'l'} f_{n'}^i(\boldsymbol{a}_{l'})\int \varphi_n^*(\boldsymbol{r}-\boldsymbol{a}_l)\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) + \mathcal{U}(\boldsymbol{r}) - \mathcal{E}_i\right]\times$$

$$\times\,\varphi_{n'}(\boldsymbol{r}-\boldsymbol{a}_{l'})\,d^3r = 0. \qquad (1.7)$$

Если поле $\mathcal{U}(\boldsymbol{r})$ меняется плавно и достаточно медленно на расстоянии постоянной решетки, так что можно положить $\mathcal{U}(\boldsymbol{r}) \approx \mathcal{U}(\boldsymbol{a}_{l'})$, то интеграл от двух последних слагаемых в

квадратной скобке равен

$$\int \varphi_n^*(\boldsymbol{r}-\boldsymbol{a}_l)\,[\mathcal{U}(\boldsymbol{r})-\mathcal{E}_i]\,\varphi_{n'}(\boldsymbol{r}-\boldsymbol{a}_{l'})\,d^3r = \\ =[\mathcal{U}(\boldsymbol{a}_{l'})-\mathcal{E}_i]\int \varphi_n^*(\boldsymbol{r}-\boldsymbol{a}_l)\,\varphi_{n'}(\boldsymbol{r}-\boldsymbol{a}_{l'})\,d^3r = \\ =[\mathcal{U}(\boldsymbol{a}_{l'})-\mathcal{E}_i]\,\delta_{nn'}\delta_{ll'}, \tag{1.8}$$

где мы воспользовались ортонормированностью функций Ванье (1.4). При подстановке (1.8) в (1.7) получим

$$\sum_{n'l'}[\mathcal{U}(\boldsymbol{a}_{l'})-\mathcal{E}_i]\,f^i_{n'}(\boldsymbol{a}_{l'})\,\delta_{nn'}\delta_{ll'} = [\mathcal{U}(\boldsymbol{a}_l)-\mathcal{E}_i]\,f^i_n(\boldsymbol{a}_l). \tag{1.9}$$

Первые два слагаемых в квадратной скобке (1.7) при переходе от функций Ванье к функциям Блоха (1.3) дают

$$\sum_{n'l'} f^i_{n'}(\boldsymbol{a}_{l'})\,\frac{1}{N}\int\left\{\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\boldsymbol{a}_l}\psi^*_{n\boldsymbol{k}}(\boldsymbol{r})\left[-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r})\right]\times\right. \\ \left.\times\sum_{\boldsymbol{k}'} e^{-i\boldsymbol{k}'\boldsymbol{a}_{l'}}\psi_{n'\boldsymbol{k}'}(\boldsymbol{r})\,d^3r\right\} = \\ =\frac{1}{N}\sum_{n'l'} f^i_{n'}(\boldsymbol{a}_{l'})\sum_{\boldsymbol{k}}\sum_{\boldsymbol{k}'} e^{i(\boldsymbol{k}\boldsymbol{a}_l-\boldsymbol{k}'\boldsymbol{a}_{l'})}\,\varepsilon_{n'}(\boldsymbol{k}')\int\psi^*_{n\boldsymbol{k}}\psi_{n'\boldsymbol{k}'}\,d^3r = \\ =\frac{1}{N}\sum_{n'l'}\sum_{\boldsymbol{k}\boldsymbol{k}'} f^i_{n'}(\boldsymbol{a}_{l'})\,e^{i(\boldsymbol{k}\boldsymbol{a}_l-\boldsymbol{k}'\boldsymbol{a}_{l'})}\,\varepsilon_{n'}(\boldsymbol{k}')\,\delta_{nn'}\delta_{\boldsymbol{k}\boldsymbol{k}'} = \\ =\frac{1}{N}\sum_{l'}\sum_{\boldsymbol{k}} f^i_n(\boldsymbol{a}_{l'})\,e^{i\boldsymbol{k}(\boldsymbol{a}_l-\boldsymbol{a}_{l'})}\,\varepsilon_n(\boldsymbol{k}), \tag{1.10}$$

где мы воспользовались ортонормированностью функций Блоха и обозначили собственное значение невозмущенного гамильтониана $-(\hbar^2/2m)\nabla^2+V(\boldsymbol{r})$ через $\varepsilon_{n'}(\boldsymbol{k}')$.

Положим $\boldsymbol{a}_l-\boldsymbol{a}_{l'}=\boldsymbol{a}_m$, тогда правая часть (1.10) равна

$$\frac{1}{N}\sum_{m}\sum_{\boldsymbol{k}} f^i_n(\boldsymbol{a}_l-\boldsymbol{a}_m)\,e^{i\boldsymbol{k}\boldsymbol{a}_m}\varepsilon_n(\boldsymbol{k}). \tag{1.11}$$

Разложим функцию $f^i_n(\boldsymbol{a}_l-\boldsymbol{r})$ в ряд Тейлора по $\boldsymbol{r}$ вблизи точки $\boldsymbol{a}_l$:

$$f^i_n(\boldsymbol{a}_l-\boldsymbol{r}) = f^i_n(\boldsymbol{a}_l)-\boldsymbol{r}\nabla f^i_n(\boldsymbol{a}_l)+{}^1\!/_2\boldsymbol{r}\boldsymbol{r}\nabla\nabla f^i_n(\boldsymbol{a}_l)-\ldots = \\ =[1-\boldsymbol{r}\nabla+{}^1\!/_2\boldsymbol{r}\boldsymbol{r}\nabla\nabla-\ldots]\,f^i_n(\boldsymbol{a}_l). \tag{1.12}$$

Здесь $\nabla f^i_n(\boldsymbol{a}_l)=[\nabla f^i_n(\boldsymbol{r})]_{\boldsymbol{r}=\boldsymbol{a}_l}$ и аналогично для выражения $\nabla\nabla f^i_n(\boldsymbol{a}_l)$ (типичный член которого имеет вид $[\partial^2 f^i_n(\boldsymbol{r})/\partial x\,\partial y]_{\boldsymbol{r}=\boldsymbol{a}_l}$).

Выражение в квадратных скобках в (1.12) можно рассматривать как оператор $\exp(-\boldsymbol{r}\nabla)$.

Подставляя в (1.12) $\boldsymbol{r}=\boldsymbol{a}_m$, получим

$$f^i_n(\boldsymbol{a}_i-\boldsymbol{a}_m) = e^{-\boldsymbol{a}_m\nabla}f^i_n(\boldsymbol{a}_l). \tag{1.13}$$

В (IV.3.10) было отмечено, что

$$\varepsilon_n(\boldsymbol{k}) = \sum_{m'} c_{m'} e^{-i\boldsymbol{k}\boldsymbol{a}_{m'}} \tag{1.14}$$

(поскольку в (IV.3.10) $\boldsymbol{l} = \{l_1, l_2, l_3\}$ принимают как положительные, так и отрицательные значения, всегда можно изменить знак у показателя экспоненты). Подставляя (1.13) и (1.14) в (1.11), получим

$$\frac{1}{N} \sum_{m} \sum_{k} \sum_{m'} e^{-i\boldsymbol{a}_m \nabla} f_n^i(\boldsymbol{a}_l)\, c_{m'} e^{i\boldsymbol{k}(\boldsymbol{a}_m - \boldsymbol{a}_{m'})} =$$

$$= \sum_{m} \sum_{m'} e^{-i\boldsymbol{a}_m \nabla} f_n^i(\boldsymbol{a}_l)\, c_{m'} \delta_{mm'} = \left[\sum_{m} c_m e^{-\boldsymbol{a}_m \nabla}\right] f_n^i(\boldsymbol{a}_l), \tag{1.15}$$

где мы воспользовались (П.6.8). Сравнивая последнее выражение с (1.14), видим, что (1.11) равно

$$\varepsilon_n(-i\nabla) f_n^i(\boldsymbol{a}_l). \tag{1.16}$$

Это выражение аналогично (IV.3.12).

Объединяя (1.16) и (1.9), получим

$$\varepsilon_n(-i\nabla) f_n^i(\boldsymbol{a}_l) + \mathcal{U}(\boldsymbol{a}_l) f_n^i(\boldsymbol{a}_l) = \mathcal{E}_i f_n^i(\boldsymbol{a}_l). \tag{1.17}$$

Для определения «плавной» функции $f_n^i(\boldsymbol{r})$, принимающей в точках $\boldsymbol{r} = \boldsymbol{a}_l$ значение $f_n^i(\boldsymbol{a}_l)$, получим дифференциальное уравнение

$$\varepsilon_n(-i\nabla) f_n^i(\boldsymbol{r}) + \mathcal{U}(\boldsymbol{r}) f_n^i(\boldsymbol{r}) = \mathcal{E}_i f_n^i(\boldsymbol{r}). \tag{1.18}$$

Это уравнение и определяет искомые коэффициенты в разложении (1.6). Оператор $\varepsilon_n(-i\nabla)$ имеет тот же смысл, что и в уравнении (IV.3.13).

Если энергия $\varepsilon_n(\boldsymbol{k})$ в точке $\boldsymbol{k} = 0$ обладает минимум (максимумом) и сферически симметрична, то в квадратичном приближении $\varepsilon_n(-i\nabla) = -(\hbar^2/2m^*)\nabla^2$ (см. (IV.3.22а)); тогда для сферически симметричного поля $\mathcal{U}(\boldsymbol{r}) = \mathcal{U}(r)$, уравнение (1.18) приобретает вид

$$-\frac{\hbar^2}{2m^*} \nabla^2 f_n^i(\boldsymbol{r}) + \mathcal{U}(\boldsymbol{r}) f_n^i(\boldsymbol{r}) = \mathcal{E}_i f_n^i(\boldsymbol{r}), \tag{1.18а}$$

т. е. является волновым уравнением для частицы с эффективной массой m^* в центральном поле $\mathcal{U}(r)$. Более общий вид рассмотрен в гл. IV, § 3, п. 3.

Если электрон притягивается к примесному центру в поле $\mathcal{U}(r)$ (донору), то он может быть захвачен на локальный уровень, лежащий в запрещенной зоне ниже зоны проводимости. Если же электрон отталкивается от примесного центра в поле $\mathcal{U}(r)$ (отрицательного иона акцептора), а следовательно, дырка притягивается к этому центру, она может быть захвачена на локальный уровень, лежащий в запрещенной зоне выше валентной зоны.

3. При выводе уравнения (1.18) было использовано предположение о медленном изменении добавочного поля $\mathcal{U}(\boldsymbol{r})$. Рассмотрим метод, который может быть применен в том случае, когда $\mathcal{U}(\boldsymbol{r})$ изменяется так быстро, что можно считать, что $\mathcal{U}(\boldsymbol{r})$ отлично от нуля только в пределах одной кристаллической ячейки. В этом случае более удобно разлагать возмущенную волновую функцию электрона $\Phi_i(\boldsymbol{r})$ по блоховским волновым функциям $\psi_k(\boldsymbol{r}) = u_k(\boldsymbol{r}) \exp(i\boldsymbol{k}\boldsymbol{r})$ (мы используем блоховские функции одной энергетической зоны):

$$\Phi_i(\boldsymbol{r}) = \sum_{\boldsymbol{k}'} f_i(\boldsymbol{k}')\, \psi_{\boldsymbol{k}'}(\boldsymbol{r}), \tag{1.19}$$

где суммирование по $\boldsymbol{k}'$ распространяется на все его квазидискретные значения в пределах 1-й бриллюэновской зоны. Подставляя выражение (1.19) в уравнение (1.5), умножая слева на $\psi^*_{\boldsymbol{k}}(\boldsymbol{r})$ и интегрируя по $\boldsymbol{r}$, получим

$$\int \psi^*_{\boldsymbol{k}}(\boldsymbol{r}) [\hat{\mathcal{H}}_0 + \mathcal{U}(\boldsymbol{r})] \sum_{\boldsymbol{k}'} f_i(\boldsymbol{k}')\, \psi_{\boldsymbol{k}'}(\boldsymbol{r})\, d\tau =$$

$$= \mathcal{E}_i \sum_{\boldsymbol{k}'} f_i(\boldsymbol{k}') \int \psi^*_{\boldsymbol{k}}(\boldsymbol{r})\, \psi_{\boldsymbol{k}'}(\boldsymbol{r})\, d\tau, \tag{1.20}$$

где $\hat{\mathcal{H}}_0 = -\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})$ — невозмущенный гамильтониан.

Так как $\hat{\mathcal{H}}_0 \psi_{\boldsymbol{k}'} = \varepsilon(\boldsymbol{k}')\, \psi_{\boldsymbol{k}'}$ и, кроме того, блоховские функции ортонормированы, то (1.20) равно

$$\varepsilon(\boldsymbol{k}) f_i(\boldsymbol{k}) + \sum_{\boldsymbol{k}'} f_i(\boldsymbol{k}') \int \psi^*_{\boldsymbol{k}}(\boldsymbol{r})\, \mathcal{U}(\boldsymbol{r})\, \psi_{\boldsymbol{k}'}(\boldsymbol{r})\, d\tau = \mathcal{E}_i f_i(\boldsymbol{k}). \tag{1.21}$$

Заменим в интеграле этого равенства блоховские функции на функции Ванье (1.2), тогда

$$\int \psi^*_{\boldsymbol{k}}(\boldsymbol{r})\, \mathcal{U}(\boldsymbol{r})\, \psi_{\boldsymbol{k}'}(\boldsymbol{r})\, d\tau =$$

$$= \frac{1}{N} \sum_{nn'} e^{i\boldsymbol{k}(\boldsymbol{a}'_n - \boldsymbol{a}_n)} \int \varphi^*(\boldsymbol{r} - \boldsymbol{a}_n)\, \mathcal{U}(\boldsymbol{r})\, \varphi(\boldsymbol{r} - \boldsymbol{a}'_n)\, d\tau. \tag{1.22}$$

Если возмущение $\mathcal{U}(\boldsymbol{r})$ отлично от нуля только в пределах одной, например нулевой, ячейки кристалла[1]), то в двойной сумме достаточно учесть один член с $\boldsymbol{n} = \boldsymbol{n}' = 0$, и мы получим вместо выражения (1.21)

$$\varepsilon(\boldsymbol{k}) f_i(\boldsymbol{k}) + \frac{\mathcal{U}_0}{N} \sum_{\boldsymbol{k}'} f_i(\boldsymbol{k}') = \mathcal{E}_i f_i(\boldsymbol{k}), \tag{1.23}$$

где $\mathcal{U}_0$ — среднее значение возмущающего потенциала $\mathcal{U}(\boldsymbol{r})$ по объему нулевой ячейки, вынесенное из-под интеграла в (1.22).

---

[1]) Очевидно, что для безграничного кристалла результат не зависит от того, в какой ячейке сосредоточено возмущение $\mathcal{U}(\boldsymbol{r})$.

Обозначая постоянную $\left(\frac{\mathcal{U}_0}{N}\right)\sum_{k'} f_i(k') = A$, получим из выражения (1.23)

$$f_i(k) = \frac{A}{\mathcal{E}_i - \varepsilon(k)}. \tag{1.24}$$

Подставляя этот результат в уравнение (1.23) и сокращая обе части равенства на неизвестную константу A, получим

$$\sum_{k'} \frac{1}{\mathcal{E}_i - \varepsilon(k')} = \frac{N}{\mathcal{U}_0} \tag{1.25}$$

— уравнение для определения собственных значений энергии $\mathcal{E}_i$ при данном возмущении $\mathcal{U}_0$. Левая часть (1.25) равна сумме дробей вида $[\mathcal{E}_i - \varepsilon(k_1)]^{-1}$, где k_1 — одно из N возможных значений волнового вектора k в первой бриллюэновской зоне. Если возмущение $\mathcal{U}_0 = 0$, то правая часть (1.25) равна бесконечности и, следовательно, собственные значения энергии электрона $\mathcal{E}_i$ для идеального кристалла совпадают с одним из N значений $\varepsilon(k)$. Можно показать, что каждому положительному и отрицательному значению $\mathcal{U}_0$ также соответствует N собственных значений $\mathcal{E}_i$ $(i = 1, 2, \ldots, N)$. До тех пор, пока $|\mathcal{U}_0|$ меньше некоторой величины, собственные значения $\mathcal{E}_i$ практически совпадают с невозмущенными собственными значениями $\varepsilon(k)$. Если же, например, $\mathcal{U}_0 < 0$ и превышает по абсолютной величине указанное выше значение, то $\mathcal{E}_i$, соответствующая нижнему краю энергетической зоны $\varepsilon(0)$, испытывает сильное возмущение — происходит отщепление уровня от дна зоны. Для $\mathcal{U}_0 > 0$ аналогичная картина имеет место для верхнего края зоны. На рис. V.1 представлена зависимость собственных значений $\mathcal{E}_i$ от величины возмущающего потенциала $\mathcal{U}_0$. Полученные результаты могут быть интерпретированы следующим образом. Электроны вблизи нижнего (верхнего) края энергетической зоны могут рассматриваться как свободные с положительной (отрицательной) массой m^*. Известно, что если глубина трехмерной потенциальной ямы $|\mathcal{U}_0|$ меньше некоторого значения, то частица не имеет в ней связанных состояний[1]). В этом случае возмущение $\mathcal{U}_0$ только рассеивает частицу, мало изменяя ее энергетический спектр.

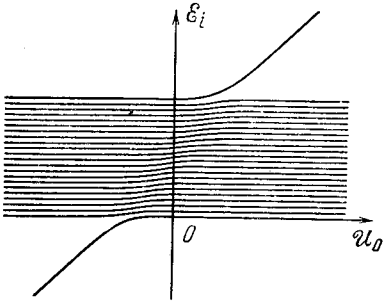


Рис. V.1.

[1]) Для сферической потенциальной ямы диаметра a это критическое значение равно $\pi^2\hbar^2/2m^*a^2$ (Шифф Л. Квантовая механика.— М.: ИЛ, 1957, с. 100).

§ 2. Локализованные состояния электрона в неидеальной решетке

1. Электрические свойства полупроводников и, в первую очередь, число носителей тока в них в ряде случаев существенно зависят от различных дефектов кристаллической решетки. К числу таких дефектов мы относим: 1) посторонние (примесные) атомы, которые могут или входить в решетку, замещая один из основных атомов, или находятся в междоузлиях; 2) основные атомы решетки, перешедшие в междоузлия; 3) пустые (вакантные) узлы решетки; 4) дислокации; 5) поверхность кристалла и т. п. Каждое такое нарушение идеальной решетки создает дополнительное поле, действующее на электрон, и может рассматриваться методами, изложенными в предыдущем параграфе.

Рассмотрим, например, атом мышьяка, имеющий пять валентных электронов, помещенный в один из узлов кристаллической решетки кремния. Четыре валентных электрона мышьяка, подобно четырем электронам атома кремния, будут участвовать в четырех направленных валентных связях решетки. Можно считать, что пятый валентный электрон, слабее связанный с узлами кристалла, будет двигаться в поле решетки и однозарядного иона мышьяка. В силу медленности изменения кулоновского потенциала задачу можно решить в приближении метода эффективной массы, рассмотренного в первом пункте предыдущего параграфа [1]). Пятый электрон атома мышьяка, не участвующий в валентных связях кристалла, но обеспечивающий его нейтральность, может находиться как в связанном состоянии вблизи иона, так и в свободном состоянии в зоне проводимости кремния. Можно пытаться грубо учесть влияние кристалла на кулоновское поле примесного иона, считая что последний погружен в среду с диэлектрической постоянной ε_0. В этом случае дополнительное поле в (1.18а) равно

$$\mathcal{U}(\boldsymbol{r}) = -e^2/\varepsilon_0 r. \tag{2.1}$$

Для простой зоны энергия электрона $\varepsilon = \hbar^2 k^2/2m^*$, и уравнение (1.18а) сводится к задаче об атоме водорода [2]) с массой электрона m^* и зарядом $e^* = \frac{e}{\sqrt{\varepsilon_0}}$. Если электрон находится в связанном состоянии вблизи примесного иона, то его энергия

$$\mathcal{E}_i = \mathcal{E}_n = -\frac{m^*(e^*)^4}{2\hbar^2 n^2} = \frac{me^4}{2\hbar^2 n^2}\frac{m^*}{m}\frac{1}{\varepsilon_0^2} = -\frac{13{,}5}{n^2}\left(\frac{m^*}{m}\right)\frac{1}{\varepsilon_0^2}\ эв, \tag{2.2}$$

---

[1]) Строго говоря, дополнительное поле $\mathcal{U}(\boldsymbol{r})$, действующее в этом случае на пятый электрон, есть разность поля иона мышьяка и поля атома кремния, однако последнее в области наиболее вероятного пребывания электрона мало.

[2]) Блохинцев Д. И., § 51.

где n — главное квантовое число и m — масса электрона в вакууме. Так как в (1.18а) предполагается, что нижнему краю зоны проводимости соответствует энергия $\mathscr{E}_0 = 0$, то отрицательные энергии электрона в связанном состоянии (2.2) расположены в запрещенной зоне. Если примесный ион расположен в l-м узле решетки, то для основного квантового состояния

$$f_0(r) = \frac{1}{\sqrt{\pi a_B^3}} e^{\frac{-|r-a_l|}{a_B}}, \qquad (2.3)$$

где боровский радиус

$$a_B = \frac{\hbar^2}{m^*(e^*)^2} = \frac{\hbar^2}{me^2}\left(\frac{m}{m^*}\right)\varepsilon_0 = 0{,}53\left(\frac{m}{m^*}\right)\varepsilon_0 \text{Å}.$$

Из (2.3)

$$f_0(a_n) = \frac{1}{\sqrt{\pi a_B^3}} e^{\frac{-|a_n-a_l|}{a_B}}. \qquad (2.3а)$$

Подставляя выражение (2.3а) в (1.6) видим, что волновая функция электрона $\Phi_0(r)$ экспоненциально затухает при удалении от l-го узла, в котором находится примесный ион. В результате возбуждения (теплового, светового и т. п.) электрон может из связанного состояния вблизи примесного иона перейти в свободное — в зону проводимости, в которой он превращается в носителя электрического тока. Такие примесные центры, которые в результате возбуждения могут поставлять свободные заряды в зону проводимости, называются *донорами*.

Представим себе теперь атом элемента третьей группы (бор, алюминий, индий или галлий) помещенным в один из пустых узлов решетки кремния или германия. У атомов 3-й группы имеются три валентных электрона в s- и p-состояниях, и для того, чтобы они образовали четырехвалентную связь в решетке Si или Ge, им необходимо заимствовать один электрон из «резервуара» валентных электронов основных атомов кристалла. Таким образом, в результате теплового или светового возбуждения электрон из валентной зоны кристалла может присоединиться к нейтральному примесному атому, образуя отрицательный ион.

Примесные центры, которые при возбуждении могут захватить электрон из валентной зоны, создавая в ней положительно заряженные дырки, называются *акцепторами*.

Таким образом, в результате возбуждения доноры превращаются в положительные ионы, создавая в зоне проводимости свободные электроны, а акцепторы — в отрицательные ионы, создавая в валентной зоне положительно заряженные дырки.

Из формулы (2.2) легко оценить, что для германия и кремния энергии основного состояния ($n = 1$) для доноров или акцепторов порядка 0,01 *эв*. Столь малые значения $|\mathscr{E}_0|$ связаны с большими

значениями диэлектрической постоянной для Ge ($\varepsilon_0 \approx 16$) и Si ($\varepsilon_0 \approx 12$). Такие уровни, получившие название *мелких*, действительно наблюдаются на опыте в германии и кремнии, когда примесными атомами являются элементы III и V групп периодической системы.

В следующей главе мы подробно рассмотрим статистическое распределение электронов между донорами, акцепторами, валентной зоной и зоной проводимости. Как мы увидим, в общем случае это распределение, а следовательно, и число носителей тока весьма чувствительно к температуре кристалла, что является характерной особенностью полупроводников.

2. При рассмотрении состояний электрона примеси в кремнии необходимо учесть тензорный характер эффективной массы и наличие шести эквивалентных минимумов энергии в зоне проводимости, расположенных на осях [100] (гл. IV, § 15); последнее приводит к шестикратному вырождению волновой функции электрона, которое частично снимается в поле кристалла.

Рассмотрим, как и в предыдущем пункте, 1*s*-состояние электрона в примесном атоме.

Мы уже отмечали, что уравнение Шредингера в приближении эффективной массы (IV.3.23), (IV.3.24) на самом деле удовлетворяется плавной частью $F(\boldsymbol{r})$ волновой функции электрона. Из-за наличия шести эквивалентных минимумов в зоне проводимости шестикратно вырожденная волновая функция электрона равна

$$\psi^{(i)}(\boldsymbol{r}) = F^{(i)}(\boldsymbol{r})\,\psi_{\boldsymbol{k}_i}(\boldsymbol{r}), \qquad i = 1,\ 2,\ \ldots,\ 6. \tag{2.4}$$

Здесь $\psi_{\boldsymbol{k}_i}(\boldsymbol{r})$ — блоховская функция электрона проводимости в точке минимума энергии $\boldsymbol{k}_i$; мы предполагаем, что при данном i функции $\psi_{\boldsymbol{k}_i}(\boldsymbol{r})$ не вырождены. Функции $F^{(i)}(\boldsymbol{r})$ удовлетворяют уравнению с эффективными массами (IV.3.23)

$$\left[-\frac{\hbar^2}{2m_\parallel}\frac{\partial^2}{\partial x_i^2} - \frac{\hbar^2}{2m_\perp}\left(\frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}\right) - \frac{e^2}{\varepsilon_0 r}\right] F^{(i)}(\boldsymbol{r}) = \mathcal{E} F^{(i)}(\boldsymbol{r}). \tag{2.5}$$

Здесь $m_\parallel = 0{,}92m$ — продольная, $m_\perp = 0{,}19m$ — поперечная эффективные массы электрона в кремнии (m — масса свободного электрона) (гл. IV, § 15). Ось x_i параллельна $\boldsymbol{k}_i$. Шесть вырожденных волновых функций (2.4) образуют базис представления (вообще говоря, приводимого) группы симметрии нашей системы, а именно: группы тетраэдра T_d. Для того чтобы определить характер расщепления шестикратно вырожденного уровня, соответствующего (2.4), необходимо определить характеры этого приводимого представления и разложить его по неприводимым представлениям группы T_d.

Для того чтобы определить характеры приводимого представления, соответствующего базису (2.4), необходимо исследовать, как преобразуются друг через друга функции $\psi_{\boldsymbol{k}_i}(\boldsymbol{r})$ при пре-

образованиях симметрии группы T_d. Очевидно, что функции $\psi_{k_i}(r)$ преобразуются друг через друга так же, как преобразуются друг через друга точки минимумов k_i. Так как точки k_i расположены на осях $\pm x$, $\pm y$, $\pm z$ на одинаковых расстояниях от начала координат, то можно сказать, что функции $\psi_{k_i}(r)$ преобразуются как шесть функций: x, $-x$, y, $-y$, z, $-z$ (если абсолютные величины всех шести координат одинаковы).

Группа T_d имеет пять классов: E, $8C_3$, $3C_4^2$, 6σ, $6S_4$ (см. табл. II.7).

Очевидно, что при единичном преобразовании E $x \to x$, $-x \to -x$, $y \to y$, ..., поэтому шпур соответствующей матрицы преобразования, т. е. характер равен 6.

Для преобразования C_3 (ось [111] имеем $x \to y \to z \to x$, $-x \to -y \to -z \to -x$, т. е. вдоль главной диагонали матрицы стоят только нули, поэтому соответствующий характер равен нулю. Для преобразования C_4^2 (ось [100] получим $x \to x$, $-x \to -x$, $y \to -y \to y$, $z \to -z \to z$. Мы видим, что на главной диагонали матрицы стоят две единицы, поэтому характер равен двум.

Совершенно аналогично можно показать, что характеры для классов σ и S_4 равны 2 и 0.

Используя табл. II.7 и полученные значения характеров, получим табл. V.1.

Таблица V.1

	E	$8C_3$	$3C_4^2$	6σ	$6S_4$
A_1	1	1	1	1	1
A_2	1	1	1	—1	—1
E	2	—1	2	0	0
F_1	3	0	—1	—1	1
F_2	3	0	—1	1	—1
$1s$	6	0	2	2	0

Используя общий прием разложения представления по неприводимым представлениям (гл. II, § 6, п. 4) или просто сравнивая характеры представления $1s$ с характерами неприводимых

представлений группы T_d, получим

$$1s = A_1 + E + F_2, \tag{2.6}$$

откуда следует, что шестикратно вырожденное состояние донора $1s$ расщепляется под действием поля кристалла на невырожденный, двукратно вырожденный и трехкратно вырожденный уровни.

3. Мы рассмотрели донорные и акцепторные состояния электронов и дырок в медленно меняющихся полях примесных ионов, когда можно пользоваться методом эффективной массы, т. е. уравнением (1.9). Как мы видели в п. 3 предыдущего параграфа, если локализованное возмущение $\mathcal{U}_0$ достигает определенной величины, то оно также приводит к отщеплению уровня энергии от дна или потолка энергетической зоны, давая связанные электронные состояния донорного или акцепторного типа.

Так как примесные ионы (в особенности отрицательные) нельзя рассматривать как точечные, то в некоторых случаях создаваемое ими поле более близко к локализованному возмущению, чем к кулоновскому потенциалу. В этом случае формула (2.2) неприменима и, как правило, появляются более глубокие примесные уровни.

Например, золото в кремнии дает донорный уровень, расположенный выше валентной зоны на 0,35 *эв*, и акцепторный уровень, расположенный на 0,54 *эв* ниже зоны проводимости [1]).

После присоединения одного электрона к положительному иону образуется нейтральный атом, поле которого с хорошим приближением описывается экранирующим потенциалом вида

$$\mathcal{U}(r) = -A\frac{e^{-\alpha r}}{r}, \tag{2.7}$$

где $1/\alpha$ — эффективный радиус действия поля. Если $1/\alpha$ порядка постоянной решетки, то возмущение (2.7) можно рассматривать как локализованное. Так как в поле потенциала (2.7) при не слишком малых $1/\alpha$ могут существовать связанные состояния электронов, то такой атом может захватить еще один электрон, поэтому в некоторых случаях примесные уровни могут быть многозарядны.

Из опыта известно, что в германии одновалентные Cu и Au могут присоединять 1, 2 и даже 3 электрона, а двухвалентные Zn и Cd могут присоединять 1 или 2 электрона.

Отметим, что количественная теория глубоких примесных уровней до настоящего времени не развита.

4. Быть может наиболее простым видом локализованного возмущения периодического поля идеального кристалла является его поверхность. Связанные состояния электронов на свободной

[1]) Ширина запрещенной зоны в кремнии при 300°K равна 1,12 *эв*.

поверхности кристалла были впервые предсказаны и теоретически рассмотрены И. Е. Таммом (1932) и получили название *таммовских поверхностных состояний*. Экспериментально их существование было подтверждено при изучении контактных явлений в полупроводниках.

Мы рассмотрим поверхностные состояния на простейшей модели одномерной ограниченной кристаллической решетки (рис. V.2), которая приближенно описывает положение в трехмерной решетке, если интересоваться только направлением, перпендикулярным к свободной поверхности кристалла. Пусть кристалл заполняет пространство при $x \geqslant 0$, где потенциальная энергия электрона $\mathcal{U}(x)$ — периодическая функция с периодом постоянной решетки a. Слева от свободной поверхности кристалла при $x \leqslant 0$ в вакууме потенциальная энергия электрона $\mathcal{U}(x)$ имеет постоянное значение $\mathcal{U}_0$.

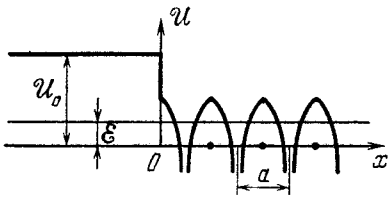


Рис. V. 2.

Уравнение Шредингера для одномерного движения электрона в поле $\mathcal{U}(x)$ имеет вид

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+\mathcal{U}(x)\,\psi=\mathcal{E}\psi. \tag{2.8}$$

Рассмотрим случай $\mathcal{E}<\mathcal{U}_0$; тогда решение (2.8) для $x \leqslant 0$ имеет вид

$$\psi_1(x)=A\exp\left(\frac{\sqrt{2m(\mathcal{U}_0-\mathcal{E})}}{\hbar}x\right), \tag{2.9}$$

в этой формуле мы отбросили слагаемое, пропорциональное $\exp\left(-\frac{\sqrt{2m(\mathcal{U}_0-\mathcal{E})}}{\hbar}x\right)$, так как оно стремится к бесконечности при $x \to -\infty$. Для $x \geqslant 0$ $\mathcal{U}(x)$ — периодическая функция и, следовательно, волновая функция имеет вид блоховской функции $u_k(x)\exp(ikx)$. Для линейного однородного дифференциального уравнения второго порядка наиболее общее решение может быть сконструировано из двух линейно независимых решений, например

$$\psi_2(x)=A_1u_k(x)\,e^{ikx}+A_2u_{-k}(x)\,e^{-ikx}. \tag{2.10}$$

Разрешенным значениям энергии электронов $\mathcal{E}(k)$ в бесконечной периодической решетке соответствуют вещественные значения волнового числа (вектора) k. В самом деле, если k комплексно, т. е. $k=k'+ik''$, то

$$\psi_2(x)=A_1u_k(x)\,e^{ik'x}e^{-k''x}+A_2u_{-k}(x)\,e^{-ik'x}e^{k''x}. \tag{2.11}$$

Совершенно ясно, что в зависимости от знака k'' либо первое, либо второе слагаемое будет стремиться к бесконечности при стремлении x к $+\infty$ или $-\infty$. Если потребовать конечности волновой функции ψ_2, то в неограниченной решетке электрон не может существовать в состоянии с комплексным k.

Иная ситуация возникает в решетке, ограниченной хотя бы с одного конца (см. рис. V.2). Если, например, $k'' > 0$, то для того, чтобы $\psi_2(x)$ оставалось конечным при $x \to +\infty$, достаточно в (2.11) положить $A_2 = 0$; для отрицательных же значений x имеет место другое решение (2.9), конечное при $x \to -\infty$.

Для того чтобы $\psi_1(x)$ и $\psi_2(x)$ описывали одно определенное состояние электрона, необходимо «сшить» решения $\psi_1(x)$ и $\psi_2(x)$ при $x = 0$, т. е.

$$\psi_1(0) = \psi_2(0), \qquad \left(\frac{d\psi_1}{dx}\right)_{x=0} = \left(\frac{d\psi_2}{dx}\right)_{x=0}. \tag{2.12}$$

Используя решения (2.9) и (2.10), получим

$$A_1 u_k(0) + A_2 u_{-k}(0) = A,$$
$$A_1 [u_k'(0) + iku_k(0)] + A_2 [u_k'(0) - iku_{-k}(0)] = A \frac{\sqrt{2m(\mathcal{U}_0 - \mathcal{E})}}{\hbar}. \tag{2.13}$$

Если k вещественно и, следовательно, $\mathcal{E}(k)$ лежит в пределах одной из разрешенных энергетических зон безграничного кристалла, то ψ_2 для $A_1 \neq 0$ и $A_2 \neq 0$ при $x \to \infty$ конечно. В этом случае два линейных уравнения (2.13) для трех неизвестных A_1, A_2 и A всегда имеют решения. Следовательно, все состояния электрона, разрешенные в безграничном кристалле, могут осуществляться и в полуограниченном.

Если же k комплексно и ему соответствует энергия $\mathcal{E}(k)$ в запрещенной зоне безграничного кристалла, то для того, чтобы ψ_2 оставалось конечным при $x \to \infty$, достаточно положить (если $k'' > 0$) $A_2 = 0$. В этом случае (2.13) превращается в однородную линейную систему двух уравнений с двумя неизвестными A_1 и A. Для того чтобы эта система имела решения, отличные от нулевого, необходимо, чтобы ее определитель равнялся нулю, что приводит к некоторому характеристическому уравнению, определяющему энергию $\mathcal{E}$. Это значение энергии может быть определено так: положим в (2.13) $A_2 = 0$, подставим $A = A_1 u_k(0)$ из первого уравнения во второе и сократим последнее на A_1, тогда

$$\mathcal{E} = \mathcal{U}_0 - \frac{\hbar^2}{2m}\left[\frac{u_k'(0)}{u_k(0)} + ik\right]^2, \tag{2.14}$$

где

$$k = k' + ik'' \ (k'' > 0).$$

Значение энергии (2.14) зависит от вида периодического потенциала $\mathcal{U}(x)$ для $x \geqslant 0$. И. Е. Тамм показал, что если периодический потенциал $\mathcal{U}(x)$ аппроксимировать прямоугольными выступами (модель Кронига и Пенни), то при определенных условиях в каждой запрещенной зоне энергии имеется один поверхностный уровень. Волновая функция электрона, находящегося на этом поверхностном уровне, экспоненциально спадает как в вакуум (2.9), так и внутрь кристалла (2.11), где $A_2 = 0$. Таким образом, электрон, находящийся на поверхностном уровне, локализован вблизи поверхности; так как он может перемещаться вдоль поверхности, то имеет место дополнительная поверхностная проводимость.

В трехмерном кристалле поверхностный уровень превращается в поверхностную зону энергии, число состояний в которой порядка 10^{15}, т. е. порядка числа атомов на 1 $см^2$ поверхности. Хотя для данного образца кристалла число поверхностных состояний может быть сравнимо с числом локальных примесных уровней, роль их обычно невелика, так как зарядка поверхности вызывает значительное изменение ее потенциала, препятствующее переходу на поверхностные уровни других электронов.

§ 3. Экситоны

1. Зонная теория, основанная на одноэлектронном приближении, учитывает все возможные квантовые состояния электрона в кристалле. При этом предполагается, что действие всех атомных ядер и всех остальных электронов на данный сводится к некоторому внешнему самосогласованному трехмерно периодическому полю. У диэлектрика или собственного полупроводника электроны полностью заполняют валентную зону, так что минимальная энергия возбуждения электрона связана с его переходом из заполненной валентной зоны в свободную зону проводимости. При каждом таком переходе возникают носители тока: электрон в зоне проводимости и дырка в валентной зоне.

Опыт, однако, показывает, что поглощение света в диэлектрике при частотах, соответствующих электронному возбуждению, не всегда сопровождается появлением носителей тока (фотопроводимостью). Как было впервые показано Я. И. Френкелем (1931), это может быть связано с особыми бестоковыми электронными возбуждениями кристалла, обладающими некоторым квазиимпульсом и энергией поступательного движения. Я. И. Френкель назвал эти возбуждения *экситонами*. Конечно, экситонные возбуждения, строго говоря, выходят за рамки одноэлектронного приближения, однако в некотором смысле их можно включить в зонную картину, если дополнительно учесть взаимодействие

между электроном в зоне проводимости и дыркой в валентной зоне (Мотт, Ванье).

В дальнейшем нас будут интересовать только связанные состояния электрона и дырки, взаимодействующие друг с другом. Если размеры такого экситона велики по сравнению с постоянной решетки, то взаимодействие электрона и дырки можно с хорошим приближением рассматривать как кулоновское взаимодействие двух точечных зарядов, ослабленное в ε_0 раз, где ε_0 — статическая диэлектрическая постоянная кристалла.

Для валентных кристаллов германия и кремния можно с достаточным для наших целей приближением положить $\varepsilon_0 = n^2$, где n — показатель преломления.

Рассмотрим уравнение Шредингера для электрона и дырки, движущихся в периодическом поле кристалла и взаимодействующих друг с другом по закону Кулона.

Пусть $\boldsymbol{r}_n$ и $\boldsymbol{r}_p$ — радиусы-векторы электрона и дырки, а $\boldsymbol{k}_n$ и $\boldsymbol{k}_p$ — волновые векторы электрона и дырки ($\boldsymbol{k}_p = -\boldsymbol{k}'_n$, где $\boldsymbol{k}'_n$ — волновой вектор электрона, соответствующего дырке в валентной зоне). Обозначим через $\varepsilon_c(\boldsymbol{k}_n)$ и $\varepsilon_v(\boldsymbol{k}_p)$ энергии электрона в зоне проводимости и дырки в валентной зоне.

Обобщая уравнение Ванье (IV.3.13) на систему, состоящую из электрона и дырки, взаимодействующих друг с другом по закону Кулона, имеем

$$\left[\varepsilon_c(-i\nabla_n) - \varepsilon_v(i\nabla_p) - \frac{e^2}{\varepsilon_0 |\boldsymbol{r}_n, \boldsymbol{r}_p|}\right] \psi(\boldsymbol{r}_n, \boldsymbol{r}_p) = \mathcal{E}\psi(\boldsymbol{r}_n, \boldsymbol{r}_p), \quad (3.1)$$

где ∇_n и ∇_p — операторы в координатах электрона и дырки, $\mathcal{E}$ — энергия всей системы; положительная кинетическая энергия ε_v отсчитывается в направлении, противоположном направлению отсчета кинетической энергии ε_c.

Если минимум энергии $\varepsilon_c(\boldsymbol{k}_n)$ и максимум энергии $\varepsilon_v(\boldsymbol{k}_p)$ расположены в центре бриллюэновской зоны и законы дисперсии энергий электронов и дырок характеризуются скалярными эффективными массами m_n и m_p, то вблизи экстремумов

$$\varepsilon_c(\boldsymbol{k}_n) = \varepsilon_c(0) + \frac{\hbar^2 k_n^2}{2m_n}, \quad \varepsilon_v(\boldsymbol{k}_p) = \varepsilon_v(0) - \frac{\hbar^2 k_p^2}{2m_p}.$$

Подставляя эти разложения в (3.1), получим [1])

$$\left[-\frac{\hbar^2}{2m_n}\nabla_n^2 - \frac{\hbar^2}{2m_p}\nabla_p^2 - \frac{e^2}{\varepsilon_0 |\boldsymbol{r}_n - \boldsymbol{r}_p|}\right] \psi(\boldsymbol{r}_n, \boldsymbol{r}_p) = (\mathcal{E} - \varepsilon_G)\psi(\boldsymbol{r}_n, \boldsymbol{r}_p), \quad (3.1a)$$

где ширина запрещенной зоны $\varepsilon_G = \varepsilon_c(0) - \varepsilon_v(0)$.

[1]) Как уже отмечалось в гл. IV, § 3 в уравнении (3.1а) стоит не полная волновая функция (3.1), а ее плавно изменяющаяся часть.

Введем радиусы-векторы $\boldsymbol{R}$ — центра тяжести электрона и дырки и $\boldsymbol{r}$ — положения электрона относительно дырки:

$$\boldsymbol{R}=\frac{m_n\boldsymbol{r}_n+m_p\boldsymbol{r}_p}{m_n+m_p}, \quad \boldsymbol{r}=\boldsymbol{r}_n-\boldsymbol{r}_p. \tag{3.2}$$

Уравнение (3.1а) в новых координатах $\boldsymbol{R}$ и $\boldsymbol{r}$ имеет вид (см. Приложение 16)

$$\left[-\frac{\hbar^2}{2M}\nabla_R^2-\frac{\hbar^2}{2\mu}\nabla_r^2-\frac{e^2}{\varepsilon_0 r}\right]\psi(\boldsymbol{R},\boldsymbol{r})=(\mathscr{E}-\varepsilon_G)\,\psi(\boldsymbol{R},\boldsymbol{r}). \tag{3.3}$$

Здесь $M=m_n+m_p$ — масса экситона, μ — приведенная масса электрона и дырки: $\frac{1}{\mu}=\frac{1}{m_n}+\frac{1}{m_p}$, ∇_R^2 и ∇_r^2 — операторы Лапласа в переменных $\boldsymbol{R}$ и $\boldsymbol{r}$.

Уравнение (3.3) решается разделением переменных

$$\psi(\boldsymbol{R},\boldsymbol{r})=\chi(\boldsymbol{R})\,\varphi(\boldsymbol{r}). \tag{3.4}$$

Подставляя (3.4) в (3.3) и деля обе части уравнения на $\chi(\boldsymbol{R})\,\varphi(\boldsymbol{r})$, получим

$$-\frac{\hbar^2}{2M}\frac{1}{\chi(\boldsymbol{R})}\nabla_R^2\chi(\boldsymbol{R})-\frac{1}{\varphi(\boldsymbol{r})}\left[\frac{\hbar^2}{2\mu}\nabla_r^2\varphi(\boldsymbol{r})+\frac{e^2}{\varepsilon_0 r}\varphi(\boldsymbol{r})\right]=\mathscr{E}-\varepsilon_G.$$

Так как первое слагаемое левой части зависит только от $\boldsymbol{R}$, второе — только от $\boldsymbol{r}$, а их сумма, равная $\mathscr{E}-\varepsilon_G$, — константа, то как первое, так и второе слагаемое — тоже постоянные величины, так что

$$-\frac{\hbar^2}{2M}\nabla_R^2\chi(\boldsymbol{R})=W\chi(\boldsymbol{R}), \tag{3.5}$$

$$-\frac{\hbar^2}{2\mu}\nabla_r^2\varphi(\boldsymbol{r})-\frac{e^2}{\varepsilon_0 r}\varphi(\boldsymbol{r})=\varepsilon\varphi(\boldsymbol{r}), \tag{3.5а}$$

где

$$W+\varepsilon=\mathscr{E}-\varepsilon_G.$$

Уравнение (3.5) описывает свободное движение частицы (экситона) с массой $M=m_n+m_p$ и энергией

$$W=\hbar^2K^2/2M, \tag{3.6}$$

где $\boldsymbol{K}$ — волновой вектор плоской волны, соответствующий свободному движению экситона.

Уравнение (3.5а) описывает относительное движение электрона и дырки, которое можно представить себе как движение электрона с массой μ вокруг неподвижной дырки. Связанным состояниям электрона и дырки соответствуют дискретные отрицательные значения энергии $\varepsilon=\varepsilon_i$; эти водородоподобные термы определяются выражением (2.2), в котором надо заменить эффективную массу m^* приведенной массой μ. Радиус экситона аналогично

боровскому радиусу атома водорода равен $\frac{\hbar^2}{me^2}\left(\frac{m}{\mu}\right)\varepsilon_0 = 0{,}53\left(\frac{m}{\mu}\right)\times$ $\times\varepsilon_0$ Å, где m — масса электрона в вакууме. Мы видим, что экситоны больших радиусов образуются в кристаллах с большой диэлектрической постоянной ε_0.

2. Экситоны малого радиуса могут быть описаны в приближении сильной связи (Я. И. Френкель, 1931). Мы воспользуемся методом Хартри — Фока. В качестве исходных одноэлектронных функций выберем волновые функции электронов в изолированных узлах решетки. Рассмотрим для простоты простую кубическую решетку из одинаковых атомов с одним электроном, и пусть основное невозбужденное состояние l-го электрона в изолированном n-м узле описывается уравнением Шредингера

$$\mathcal{H}_n\psi_n(\boldsymbol{r}_l) = \varepsilon_0\psi_n(\boldsymbol{r}_l). \tag{3.7}$$

Здесь

$$\mathcal{H}_n = -\frac{\hbar^2}{2m}\nabla_l^2 + \mathcal{U}_n(\boldsymbol{r}_l), \tag{3.7а}$$

где ∇_l^2 — оператор Лапласа для l-го электрона, а $\mathcal{U}_n(\boldsymbol{r}_l)$ — его потенциальная энергия в n-м изолированном узле решетки и ε_0 — энергия основного состояния.

Пренебрежем для простоты перекрытием волновых функций даже соседних узлов, тогда аналогично (IV.2.7)

$$\int \psi_n^*(l)\,\psi_{n'}(l)\,d\tau_l = \delta_{nn'}, \tag{3.8}$$

где l стоит вместо $\boldsymbol{r}_l$ и $d\tau_l = dx_l\,dy_l\,dz_l$. Предположим, что спины всех электронов параллельны, что, очевидно, не существенно, если мы не интересуемся магнитными свойствами системы, вопросами ковалентной связи в решетке и т. п.

Если не учитывать спин-орбитального взаимодействия, так что полная волновая функция может быть представлена в виде произведения координатной части на спиновую, то в силу симметрии последней в нашем случае координатная часть волновой функции должна быть антисимметрична. Правильная антисимметричная волновая функция всего кристалла, построенная на одноэлектронных функциях $\psi_n(l)$, имеет вид определителя (IV.2.8б):

$$\Phi(1, 2, \ldots, N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(1)\ \psi_1(2)\ldots\psi_1(N) \\ \ldots\ldots\ldots\ldots \\ \psi_N(1)\ \psi_N(2)\ldots\psi_N(N) \end{vmatrix}, \tag{3.9}$$

если в основной области кристалла имеется N узлов.

Используя (3.8), легко показать, аналогично (IV.2.8а), что

$$\int \Phi^*(1, 2, \ldots, N)\,\Phi(1, 2, \ldots, N)\,d\tau_1\,d\tau_2\ldots d\tau_N = \int \Phi^*\Phi\,d\tau = 1 \tag{3.9а}$$

$$(d\tau \equiv d\tau_1\,d\tau_2\ldots d\tau_N).$$

Гамильтониан системы всех N электронов

$$\hat{\mathcal{H}} = \sum_{n=1}^{N} \hat{\mathcal{H}}_n + V(1, 2, \ldots, N), \tag{3.10}$$

где $\sum_{n=1}^{N} \hat{\mathcal{H}}_n = \hat{\mathcal{H}}_0$ — невозмущенный гамильтониан (можно, например, считать $\hat{\mathcal{H}}_n = \frac{-\hbar^2}{2m}\nabla_n^2 + \mathcal{U}_n(r_n)$), а $V(1, 2, \ldots, N)$ включает взаимодействие всех электронов между собою и энергию каждого n-го электрона в поле всех других m-х узлов решетки ($m \neq n$). В нулевом приближении энергия системы

$$\mathcal{E}_0 = \int \Phi^* \hat{\mathcal{H}}_0 \Phi \, d\tau = N\varepsilon_0, \tag{3.11}$$

что непосредственно следует из выражений (3.9), (3.7) и (3.8).

Поправка к энергии основного состояния в первом приближении теории возмущений равна

$$\mathcal{E}_1 = \int \Phi^* V \Phi \, d\tau. \tag{3.12}$$

Обозначим волновую функцию возбужденного состояния l-го электрона у n-го узла через ψ'_n; тогда

$$\hat{\mathcal{H}}_n \psi'_n = \varepsilon_1 \psi'_n, \tag{3.13}$$

где ε_1 — энергия возбужденного состояния электрона в изолированном узле.

Будем считать, что возбужденные волновые функции данного узла также не перекрываются с волновыми функциями соседних узлов; кроме того, для данного n

$$\int \psi_n^* \psi'_n \, d\tau = 0, \tag{3.14}$$

так как возбужденное состояние ψ'_n ортогонально к основному ψ_n.

Волновая функция электронов в кристалле, у которого возбужден n-й узел, обладающая правильными антисимметричными свойствами, имеет вид

$$\Phi_n(1, 2, \ldots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \ldots & \psi_1(N) \\ \ldots & \ldots & \ldots \\ \psi'_n(1) & \ldots & \psi'_n(N) \\ \ldots & \ldots & \ldots \\ \psi_N(1) & \ldots & \psi_N(N) \end{vmatrix}. \tag{3.15}$$

Используя (3.8), в котором для $n \neq n'$ волновые функции могут относиться и к возбужденным состояниям, и (3.14), легко показать, что

$$\int \Phi_m^* \Phi_n \, d\tau_1 \ldots d\tau_N = \delta_{mn}. \tag{3.16}$$

Выражение (3.15), в противоположность (3.9), не является еще правильной волновой функцией нулевого приближения для возбужденного кристалла. В самом деле, в силу трансляционной симметрии возбуждение может быть локализовано на любом узле решетки, при этом энергия кристалла остается неизменной. Таким образом, состояние является N-кратно вырожденным и правильная волновая функция нулевого приближения должна быть построена в виде линейной суперпозиции:

$$\Phi'(1, 2, \ldots, N) = \sum_{n=1}^{N} C_n \Phi_n \quad (1, 2, \ldots, N), \tag{3.17}$$

где C_n — постоянные коэффициенты, которые должны быть определены по общим правилам теории возмущений вырожденных состояний (ср. с гл. IV, § 7, п. 2).

Вычислим действие невозмущенного оператора $\hat{\mathcal{H}}_0 = \sum_{l=1}^{N} \hat{\mathcal{H}}_l(l)$ на функцию Φ_n (3.15):

$$\hat{\mathcal{H}}_0 \Phi_n = [\hat{\mathcal{H}}_1(1) + \hat{\mathcal{H}}_2(2) + \ldots + \hat{\mathcal{H}}_l(l) + \ldots + \hat{\mathcal{H}}_N(N)] \times \\ \times \frac{1}{\sqrt{N!}} \sum_{p} (-1)^{[p]} \hat{P} \{\psi_1(p_1) \ldots \psi'_n(p_n) \ldots \psi_N(p_N)\}, \tag{3.18}$$

где $p_1, p_2, \ldots, p_n, \ldots, p_N$ — числа 1, 2, 3, ..., N, расставленные в некотором порядке. Определитель (3.15) записан здесь в виде суммы, в которой суммирование распространяется на всевозможные перестановки чисел $p_1, p_2, \ldots, p_n, \ldots, p_N$, а $[p]$ обозначает число беспорядков в этой перестановке [1]).

Следует различать два случая действия одного из операторов $\hat{\mathcal{H}}_l(l)$ ($l = 1, 2, 3, \ldots, N$) на произведение одноэлектронных функций $\{\psi_1(p_1) \ldots \psi'_n(p_n) \ldots \psi_N(p_N)\}$: 1) $p_n = l$; тогда действие оператора, согласно (3.13), сводится к умножению $\{\ \}$ на ε_1; 2) $p_n \neq l$; тогда действие оператора, согласно (3.7), сводится к умножению $\{\ \}$ на ε_0. Так как второй случай для каждого l может реализоваться $N-1$ способом, то

$$\hat{\mathcal{H}}_0 \Phi_n = [(N-1)\varepsilon_0 + \varepsilon_1] \Phi_n. \tag{3.18а}$$

Энергия кристалла в этом приближении равна

$$\int \Phi'^* \hat{\mathcal{H}}_0 \Phi' \, d\tau = \sum_m \sum_n C_m^* C_n \int \Phi_m^* \hat{\mathcal{H}}_0 \Phi_n \, d\tau = \\ = [(N-1)\varepsilon_0 + \varepsilon_1] \sum_m \sum_n C_m^* C_n \delta_{mn} = [(N-1)\varepsilon_0 + \varepsilon_1] \sum_n |C_n|^2 = \\ = (N-1)\varepsilon_0 + \varepsilon_1, \tag{3.19}$$

если нормировать на единицу волновую функцию Φ', т. е. положить

$$\sum_n |C_n|^2 = 1. \tag{3.19а}$$

[1]) Смирнов В. И. — т. 3, ч. 1, § 1.

Гамильтониану (3.10) соответствует уравнение Шредингера

$$[\mathscr{H}_0 + V(1, 2, \ldots, N)]\,\psi = W\psi. \tag{3.20}$$

Подставим сюда вместо ψ приближенную волновую функцию Φ' (3.17). Учитывая (3.18а), получим

$$\sum_n C_n \{[(N-1)\,\varepsilon_0 + \varepsilon_1]\,\Phi_n + V\Phi_n\} = W \sum_n C_n \Phi_n.$$

Умножая слева обе части равенства на Φ_m^*, интегрируя по координатам всех электронов и учитывая (3.16), получим

$$\sum_{n=1}^{N} C_n V_{mn} - \mathscr{E}' C_m = 0, \tag{3.21}$$

где

$$V_{mn} = \int \Phi_m^* V \Phi_n \, d\tau, \tag{3.21а}$$

$$\mathscr{E}' = W - [(N-1)\,\varepsilon_0 + \varepsilon_1]. \tag{3.21б}$$

Линейная однородная алгебраическая система N уравнений для N неизвестных коэффициентов C_n (3.21) в развернутом виде имеет вид:

$$\begin{aligned}
&(V_{11} - \mathscr{E}')\,C_1 + V_{12}C_2 + \ldots + V_{1N}C_N = 0,\\
&V_{21}C_1 + (V_{22} - \mathscr{E}')\,C_2 + \ldots + V_{2N}C_N = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&V_{N1}C_1 + V_{N2}C_2 + \ldots + (V_{NN} - \mathscr{E}')\,C_N = 0.
\end{aligned} \tag{3.21в}$$

Для того чтобы эта однородная система имела решения, отличные от нулевых, необходимо положить ее определитель равным нулю:

$$\begin{vmatrix}
(V_{11} - \mathscr{E}'), & V_{12} & \ldots & V_{1N}\\
V_{21}, & (V_{22} - \mathscr{E}') & \ldots & V_{2N}\\
\ldots\ldots\ldots\ldots & \ldots\ldots\ldots & \ldots & \ldots\ldots\ldots\\
V_{N1}, & V_{N2} & \ldots & (V_{NN} - \mathscr{E}')
\end{vmatrix} = 0. \tag{3.22}$$

Это характеристическое уравнение N-й степени относительно $\mathscr{E}'$ определяет в первом приближении спектр экситонного возбуждения кристалла. Мы определим спектр экситона, воспользовавшись следующим не строгим, но наглядным приемом. Аналогично случаю сильной связи в одноэлектронном приближении (гл. IV, § 7), коэффициенты C_n, учитывающие трансляционное вырождение, можно положить

$$C_n = C_0 e^{i\boldsymbol{k}\boldsymbol{a}_n}, \tag{3.23}$$

где $\boldsymbol{k}$ — волновой вектор квазичастицы — экситона.

В этом случае из уравнений (3.21) следует

$$\mathscr{E}' = \sum_n \frac{C_n}{C_m} V_{mn} = \sum_n e^{i\boldsymbol{k}(\boldsymbol{a}_n - \boldsymbol{a}_m)} V_{mn}. \tag{3.24}$$

Так как матричные элементы V_{mn}, аналогично стоящим перед ними множителям, зависят только от разности $(\boldsymbol{a}_n - \boldsymbol{a}_m)$, то можно без ограничения общности положить $\boldsymbol{a}_m = 0$, так что

$$\mathscr{E}' = \sum_n e^{i\boldsymbol{k}\boldsymbol{a}_n} V_{0n}. \tag{3.24а}$$

Учитывая в простой кубической решетке только шесть ближайших соседей ($n = 1, 2, \ldots, 6$) и полагая для них одинаковые $V_{0n} = -V_0$, получим аналогично (IV.7.10а)

$$\mathscr{E}' = -2V_0 (\cos ak_x + \cos ak_y + \cos ak_z). \tag{3.24б}$$

Таким образом, в приближении сильной связи экситонное возбуждение в простом кубическом кристалле определяется энергетической зоной (3.24б), подобной (IV.7.10а).

Аналогично зависимости (IV.7.12) и (IV.7.13) мы можем определить эффективную массу и скорость экситона:

$$m^*_{\text{экс}} = \hbar^2 / 2V_0 a^2, \tag{3.25}$$

$$\boldsymbol{v}_{\text{экс}} = \frac{2V_0 a}{\hbar} [\sin ak_x \boldsymbol{i}_0 + \sin ak_y \boldsymbol{j}_0 + \sin ak_z \boldsymbol{k}_0].$$

Волновая функция экситона

$$\Phi' (1, 2, \ldots, N) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{i\boldsymbol{k}\boldsymbol{a}_n} \Phi_n (1, 2, \ldots, N), \tag{3.26}$$

где из условий нормировки (3.19а) мы положили $C_0 = 1/\sqrt{N}$. Это выражение напоминает одноэлектронную волновую функцию в приближении сильной связи (IV.7.2а).

Величина V_0 того же порядка, что и A, входящая в выражение (IV.7.12), поэтому эффективная масса экситона, вообще говоря, того же порядка, что и эффективная масса электрона и дырки.

§ 4. Поляроны

1. В ионных кристаллах могут реализоваться особые состояния электронов проводимости, впервые подробно изученные С. И. Пекаром (1946), которые он назвал *поляронами*.

Поляроны возникают в результате поляризации ионной решетки электроном проводимости. Эта поляризация кристалла вызывает понижение энергии электрона, т. е. приводит к возникновению в области нахождения электрона потенциальной ямы. Состояние электрона проводимости, локализованного в этой потенциальной яме, описывается затухающей волновой функцией. Таким образом, возникает самосогласованное состояние: локализация электрона вызывает поляризацию кристалла, а последняя поддерживает локализацию электрона. Конечно, это автолока-

лизованное состояние электрона может свободно перемещаться по всему кристаллу.

Если размеры ψ-облака электрона в поляронном состоянии велики по сравнению с постоянной решетки (поляроны большого радиуса), то ионный кристалл можно описывать как непрерывную диэлектрическую среду (континуум).

Необходимо иметь в виду, что в образовании поляронного состояния участвует не полная поляризация кристалла, а только ее инерционная часть. В самом деле, поляризация электронных оболочек ионов, безынерционно следующая за движением электрона проводимости, входит в самосогласованный периодический потенциал, действующий на электрон [1]). Потенциальная яма поляронного состояния обусловливается только инерционной частью поляризуемости — смещением тяжелых ионов.

2. Рассмотрим так называемые *поляроны сильной связи* [2]). Самосогласованное состояние электрона проводимости в поляронном состоянии определяется из следующего уравнения Шредингера:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r})+[V(\boldsymbol{r})+\mathcal{U}(\boldsymbol{r})]\,\psi(\boldsymbol{r})=\mathcal{E}\psi(\boldsymbol{r}). \tag{4.1}$$

Здесь $V(\boldsymbol{r})$ — периодический потенциал кристалла, $\mathcal{U}(r)$ — энергия электрона в поле поляризованного им самим кристалла. Для поляронов большого радиуса можно опустить периодический потенциал $V(\boldsymbol{r})$, заменив одновременно массу свободного электрона m на его эффективную массу m^*; тогда

$$-\frac{\hbar^2}{2m^*}\nabla^2\psi+\mathcal{U}(\boldsymbol{r})\,\psi=\mathcal{E}\psi. \tag{4.2}$$

Уравнение (4.2) может быть получено из вариации функционала

$$\mathcal{E}[\psi]=\frac{\hbar^2}{2m^*}\int(\nabla\psi)^2\,d\tau+\int\mathcal{U}\psi^2\,d\tau \tag{4.3}$$

при дополнительном условии нормировки [3])

$$\int\psi^2\,d\tau=1. \tag{4.4}$$

В самом деле, вычтем из (4.3) выражение (4.4), помноженное на неопределенный множитель Лагранжа λ. Варьируя полученное выражение по ψ и приравнивая нулю, получим

$$\frac{\hbar^2}{2m^*}\int 2(\nabla\psi\nabla\,\delta\psi)\,d\tau+\int\mathcal{U}2\psi\,\delta\psi\,d\tau-\int\lambda 2\psi\,\delta\psi\,d\tau=0,$$

где было использовано, что $\delta\nabla\psi=\nabla\,\delta\psi$.

---

[1]) Пекар С. И. Исследования по электронной теории кристаллов. — М.: Гостехиздат, 1951.

[2]) См. там же, а также сб. Поляроны./Под ред. Ю. А. Фирсова. — М., 1975.

[3]) Мы рассматриваем только основное состояние полярона, для которого волновая функция ψ вещественна, поэтому $|\psi|^2=\psi^2$.

По теореме Грина

$$\int (\nabla\psi\nabla\,\delta\psi)\,d\tau = -\int \delta\psi\,\nabla^2\psi\,d\tau + \oint \delta\psi\,\nabla\psi\,d\boldsymbol{S}.$$

Последний интеграл по поверхности исчезает при стремлении поверхности $S\to\infty$, в силу быстрого убывания ψ на бесконечности.

Таким образом,

$$\int\left[-\frac{\hbar^2}{2m^*}\nabla^2\psi + \mathcal{U}\psi - \lambda\psi\right]\delta\psi\,d\tau = 0,$$

что при произвольности $\delta\psi$ и отождествлении λ с $\mathscr{E}$ совпадает с выражением (4.2).

В некоторых случаях для приближенного решения уравнения Шредингера (4.2) удобнее пользоваться вариационным принципом, т. е. функционалом $\mathscr{E}[\psi]$ (4.3). Это может быть сделано в тех случаях, когда нам из физических соображений известен общий вид волновой функции. Задавая волновую функцию в виде $\psi(\boldsymbol{r}; a_1, a_2, \ldots)$, где $a_1, a_2, \ldots$ — неопределенные пока параметры, можно вычислить посредством (4.3) энергию $\mathscr{E}$ как функцию $a_1, a_2, \ldots$ Находя минимум $\mathscr{E}(a_1, a_2, \ldots)$ в зависимости от параметров $a_1, a_2, \ldots$, мы определим наилучшее собственное значение энергии, соответствующее выбранному виду волновой функции. При этом, даже при весьма грубой аппроксимации волновой функции, собственные значения находятся часто довольно точно.

Усредненная энергия взаимодействия электрона с поляризованным кристаллом, входящая в (4.3),

$$\mathcal{U}_{\text{ср}} = \int \mathcal{U}\psi^2\,d\tau \tag{4.5}$$

может быть для диэлектрического континуума определена следующим образом.

Вектор электрической индукции, обусловленный распределенным зарядом электрона $e\,|\psi(\boldsymbol{r}')|^2\,d\tau'$, равен

$$\boldsymbol{D}(\boldsymbol{r}) = e\int |\psi(\boldsymbol{r}')|^2\,\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\,d\tau' \tag{4.6}$$

и, конечно, не зависит в однородном диэлектрике от диэлектрической постоянной. Энергия диполя с моментом $\boldsymbol{p}$ в поле $\boldsymbol{E}$, как известно, равна $-(\boldsymbol{pE})$. Энергия диполя $\boldsymbol{P}(\boldsymbol{r})\,d\tau$, где $\boldsymbol{P}$ — вектор поляризации кристалла, в электрическом поле заряда $e\,|\psi(\boldsymbol{r}')|^2\,d\tau'$ равна

$$-\boldsymbol{P}(\boldsymbol{r})\,d\tau e\,|\psi(\boldsymbol{r}')|^2\,d\tau'\,\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|^3}. \tag{4.7}$$

Интегрируя (4.7) по всем элементам поляризованного кристалла $d\tau$ и всем элементам распределенного заряда $d\tau'$, получим

$$\mathcal{U}_{\text{ср}} = -e\iint \boldsymbol{P}(\boldsymbol{r})\,|\psi(\boldsymbol{r}')|^2\,\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\,d\tau\,d\tau' = -\int \boldsymbol{D}(\boldsymbol{r})\,\boldsymbol{P}(\boldsymbol{r})\,d\tau. \tag{4.8}$$

Появление в этом выражении $\boldsymbol{D}$ вместо напряженности поля $\boldsymbol{E}$ связано с тем, что (4.7) есть энергия диполя $\boldsymbol{P}\,d\tau$ в поле заряда $e|\psi|^2 d\tau'$ в пустоте, независимо от поля, создаваемого всеми другими свободными и связанными зарядами в кристалле.

Подставляя выражение (4.8) в (4.3), получим

$$\mathscr{E}[\psi]=\frac{\hbar^2}{2m^*}\int(\nabla\psi)^2\,d\tau-\int(\boldsymbol{DP})\,d\tau. \tag{4.9}$$

Мы должны теперь учесть, что $\boldsymbol{P}$ является не полной поляризацией $\boldsymbol{P}_0$, а только ее инерционной частью.

Если ε — полная диэлектрическая постоянная ионного кристалла, то

$$\boldsymbol{P}_0=\frac{\varepsilon-1}{4\pi\varepsilon}\boldsymbol{D}. \tag{4.10}$$

Поляризация $\boldsymbol{P}_e$, связанная с деформацией электронных оболочек ионов, когда они находятся в равновесных положениях, безынерционно следующая за движением электрона проводимости, равна

$$\boldsymbol{P}_e=\frac{n^2-1}{4\pi n^2}\boldsymbol{D}, \tag{4.11}$$

где n — коэффициент преломления в длинноволновой области.

Таким образом,

$$\boldsymbol{P}=\boldsymbol{P}_0-\boldsymbol{P}_e=\frac{1}{4\pi\varepsilon^*}\boldsymbol{D}, \tag{4.12}$$

где

$$\frac{1}{\varepsilon^*}=\frac{1}{n^2}-\frac{1}{\varepsilon}. \tag{4.12a}$$

Величина ε^* играет роль эффективной диэлектрической постоянной, входящей в теорию поляронов.

Заметим, что в соответствии с вариационным принципом минимум функционала $\mathscr{E}[\psi]$ (4.3) надо искать, варьируя ψ в постоянном поле $\mathscr{U}$, т. е. при $P=\mathrm{const}$. После этого $\boldsymbol{P}$ определяется из уравнения (4.12) через соответствующее $\boldsymbol{D}[\psi]$ (4.6). Легко показать, что тот же результат получится, если заменить в выражении (4.9) $\boldsymbol{P}$ согласно (4.12), но приписать интегралу множитель $^1/_2$. Мы получим вместо (4.9) функционал

$$J[\psi]=\frac{\hbar^2}{2m^*}\int(\nabla\psi)^2\,d\tau-\frac{1}{8\pi\varepsilon^*}\int\boldsymbol{D}^2[\psi]\,d\tau. \tag{4.13}$$

Отсюда совместно с условием нормировки (4.4) может быть определена волновая функция основного состояния полярона ψ_0 и соответствующее ей собственное значение энергии $\mathscr{E}_0\equiv\mathscr{E}[\psi_0]$. В качестве пробной функции основного 1s-состояния полярона Пекар берет выражение

$$\psi=A(1+\alpha r+\beta r^2)\,e^{-\alpha r}, \tag{4.14}$$

где α и β — вариационные параметры, а A — нормировочная константа, которую можно выразить через них.

Подставляя выражение (4.14) в (4.13), вычисляя интегралы и минимизируя полученное выражение по α и β, получим

$$\psi_0 = 0{,}12\alpha_0^{3/2}(1+\alpha_0 r + 0{,}45\alpha_0^2 r^2)\, e^{-\alpha_0 r}, \tag{4.15}$$

где

$$\alpha_0 \equiv \frac{1}{r_0} = 0{,}66\,\frac{m^* e^2}{\hbar^2 \varepsilon^*} \tag{4.15а}$$

— обратный радиус полярона [1]).

Из уравнений (4.9), (4.12) и (4.6) получим для собственного значения энергии основного состояния полярона выражения

$$\mathcal{E}_0 \equiv \mathcal{E}[\psi_0] = -0{,}164\,\frac{m^* e^4}{\hbar^2 \varepsilon^{*2}}, \tag{4.16}$$

что может быть сравнено с основным термом атома водорода $-0{,}5$, $(me^4/\hbar^2)$, где m — масса свободного электрона. Обычно $m^* \leqslant m$, и так как $\varepsilon^* > 1$, то $\mathcal{E}_0$ значительно меньше энергии ионизации атома водорода (13,5 *эв*).

Используя выражение (4.15) для ψ_0 и (4.6) для $\boldsymbol{D}[\psi]$, можно показать, что [2])

$$\mathcal{U}_{\text{ср}}[\psi_0] = \mathcal{U}^0_{\text{ср}} = {}^4/_3\,\mathcal{E}_0, \tag{4.17}$$

$$\mathcal{T}[\psi_0] = {}^1/_3\,\mathcal{E}_0; \tag{4.18}$$

как следует из (4.3) и (4.13),

$$\mathcal{E}[\psi] = \mathcal{T} + \mathcal{U}_{\text{ср}}, \tag{4.19}$$

$$\mathcal{T}[\psi] = \mathcal{T} + {}^1/_2\,\mathcal{U}_{\text{ср}}, \tag{4.20}$$

где $\mathcal{T} \equiv \left(\frac{1}{2m^*}\right)\int (\nabla\psi)^2\, d\tau$ — кинетическая энергия электрона.

С другой стороны, энергия инерционно поляризованного кристалла равна

$$\mathcal{U}_p = \int \frac{\boldsymbol{E}\boldsymbol{D}}{8\pi}\, d\tau = \frac{1}{8\pi\varepsilon^*}\int \boldsymbol{D}^2\, d\tau, \tag{4.21}$$

что для основного состояния дает

$$\mathcal{U}_p[\psi_0] = \mathcal{U}^0_p = -{}^2/_3\,\mathcal{E}_0. \tag{4.21а}$$

При тепловой диссоциации полярона последний распадается в результате тепловой флуктуации поляризационной ямы, поэтому энергия активации равна сумме $-\mathcal{E}_0$ и $-\mathcal{U}^0_p$, т. е.

$$W_{\text{т}} = -\mathcal{E}_0 - \mathcal{U}^0_p = -\mathcal{E}_0 + {}^2/_3\,\mathcal{E}_0 = -{}^1/_3\,\mathcal{E}_0 = -\mathcal{T}[\psi_0]. \tag{4.22}$$

---

[1]) В дальнейшем индексом нуль мы будем отмечать величины, относящиеся к основному состоянию полярона.

[2]) Пекар С. И. Исследования по электронной теории кристаллов.— М.: Гостехиздат, 1951.

Красная граница фотодиссоциации полярона определяется энергией, необходимой для перехода электрона из основного состояния $-\mathscr{E}_0$ в нижний край зоны проводимости при неизменной конфигурации ионов (принцип Франка—Кондона), т. е.

$$W_{ph} = -\mathscr{E}_0 = 3W_т. \tag{4.22а}$$

Подчеркнем, что выведенные выше соотношения

$$|W_т| : |\mathcal{U}_p| : |W_{ph}| : |\mathcal{U}_{ср}| = 1:2:3:4$$

имеют место только для поляронов сильной связи.

Основное состояние полярона сильной связи может быть приближенно определено следующим простым способом. Минимальная кинетическая энергия электрона с массой m^* в сферической полости радиуса r с непроницаемыми стенками равна[1])

$$\mathscr{T} = \frac{\pi^2\hbar^2}{2m^*r^2}. \tag{4.23}$$

Потенциальная энергия взаимодействия этого электрона с инерционно поляризующимся континуумом, порядка

$$\mathcal{U}_{ср} \approx -\frac{e^2}{\varepsilon^* r}. \tag{4.24}$$

Минимум полной энергии $\mathscr{E} = \mathscr{T} + \mathcal{U}_{ср}$ в зависимости от r равен

$$\mathscr{E}_0 = -\frac{1}{\pi^2}\frac{m^*e^4}{2\varepsilon^{*2}\hbar^2} = -0{,}051\,\frac{m^*e^4}{\varepsilon^{*2}\hbar^2} \tag{4.25}$$

при

$$r_0 = \pi^2\frac{\varepsilon^*\hbar^2}{m^*e^2}. \tag{4.26}$$

Сравнивая выражения (4.25) и (4.26) с выражениями (4.16) и (4.15а) видим, что приближенный расчет практически приводит к тем же формулам, что и вариационный метод (отличие только в численных множителях).

Для того чтобы реализовались поляроны сильной связи, необходимо чтобы движение электрона в поляронной яме было быстрым по сравнению с движением тяжелых ионов (адиабатическое приближение), т. е.

$$\frac{r_0}{v} \ll \frac{1}{\omega}, \tag{4.27}$$

где v—скорость электрона, ω—предельная частота колебаний ионов оптической ветви. При этом для электрона в основном состоянии длина волны де Бройля ($\lambda\!\!\!^{-} = \lambda/2\pi$) равна

$$\lambda\!\!\!^{-} \approx r_0 \ll \frac{v}{\omega} \equiv l, \tag{4.28}$$

[1]) См. ссылку на стр. 310.

где l — путь, проходимый электроном за время одного колебания ионов. Кинетическая энергия электрона

$$\mathscr{T}_0 = {}^1/_2\, m^* v^2. \tag{4.29}$$

Исключая из неравенства (4.27) v и r_0 посредством (4.29), (4.23) (для $r = r_0$) и (4.26), получим

$$\frac{1}{\pi^3}\frac{m^* e^4}{\varepsilon^{*2}\hbar^2}\frac{1}{\hbar\omega} = \frac{2}{\pi}\frac{|\mathscr{E}_0|}{\hbar\omega} \gg 1, \tag{4.30}$$

что совпадает с обычным условием применимости адиабатического приближения (энергия быстрой подсистемы много больше энергии медленной подсистемы).

Вводя безразмерный параметр

$$\alpha^2 = \frac{\pi^2 |\mathscr{E}_0|}{\hbar\omega} = \frac{m^* e^4}{2\varepsilon^{*2}\hbar^3\omega}, \tag{4.31}$$

получим для энергии основного состояния поляронов сильной связи

$$\mathscr{E}_0 = -\frac{1}{\pi^2}\alpha^2\hbar\omega \approx -\frac{1}{10}\alpha^2\hbar\omega \tag{4.32}$$

при условии

$$\alpha \gg 1. \tag{4.33}$$

Таким образом, в этом случае энергия пропорциональна α^2.

3. Рассмотрим теперь другой предельный случай — поляроны слабой связи (Г. Фрелих, Г. Пельцер, С. Зинау, 1950), когда параметр

$$\alpha \ll 1. \tag{4.34}$$

Определим в этом случае влияние взаимодействия электрона с поляризованным им ионным кристаллом на энергетический спектр электрона.

Рассмотрим длинные оптические волны в бинарном ионном кристалле в континуальном приближении (гл. III, § 9, п. 2). Из (III.9.37) видно, что электрическое поле $\boldsymbol{E}$ в кристалле вызывается только продольными колебаниями $\boldsymbol{w}_l$; это представляется естественным, так как только при продольных колебаниях происходит изменение объема кристалла, в результате чего возникает в нем связанный электрический заряд.

Подставляя $\boldsymbol{E}$ из (III.9.37) в (III.9.32), получим для вектора поляризации

$$\boldsymbol{P} = \omega_0\sqrt{\frac{\varepsilon_0-\varepsilon_\infty}{4\pi}}\,\boldsymbol{w} - \frac{\varepsilon_\infty - 1}{4\pi\varepsilon_\infty}\omega_0\sqrt{4\pi(\varepsilon_0-\varepsilon_\infty)}\,\boldsymbol{w}_l = \sqrt{\frac{N_0 m_r \omega_l^2}{4\pi\varepsilon^*}}\,(\boldsymbol{u}_+ - \boldsymbol{u}_-). \tag{4.35}$$

Как мы увидим ниже, с электроном взаимодействуют только продольные колебания, поэтому в (4.35) на самом деле $\boldsymbol{w}_l = \boldsymbol{w}$. Здесь N_0 — число кристаллических ячеек в единице объема, приведенная масса ионов $m_r = m_+ m_- / (m_+ + m_-)$, $\omega_l = (\varepsilon_0/\varepsilon_\infty)^{1/2}\omega_0$ — предельная частота продольных оптических колебаний, $\frac{1}{\varepsilon^*} = \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0}$, где ε_∞ и ε_0 — высокочастотная и статическая диэлектрические постоянные, $\boldsymbol{w} = \sqrt{N_0 m_r}\,\boldsymbol{s} = \sqrt{N_0 m_r}\,(\boldsymbol{u}_+ - \boldsymbol{u}_-)$, где $\boldsymbol{s} = \boldsymbol{u}_+ - \boldsymbol{u}_-$ — смещение положительного иона относительно отрицательного. Считая для положительного и отрицательного ионов $k = 1$ и $k = 2$, имеем из (III.6.4) в континуальном приближении ($\boldsymbol{a}_n = \boldsymbol{r}$)

$$\boldsymbol{u}^k(\boldsymbol{r}) = \frac{1}{\sqrt{N_0 m_k}} \sum_{q,j} \boldsymbol{e}_{jk}(\boldsymbol{q})\, a_j(\boldsymbol{q},\ t)\, e^{i\boldsymbol{q}\boldsymbol{r}}. \tag{4.36}$$

Для длинных оптических волн ($\boldsymbol{q} \to 0$) векторы поляризации $\boldsymbol{e}_{jk}$ вещественны и ионы колеблются относительно неподвижного центра тяжести ячейки, т. е.

$$\sqrt{m_1}\,\boldsymbol{e}_{j1} + \sqrt{m_2}\,\boldsymbol{e}_{j2} = 0, \tag{4.37}$$

как это следует из (III.5.22).

С другой стороны, условие нормировки (III.6.2) имеет вид

$$\boldsymbol{e}_{j1}^2 + \boldsymbol{e}_{j2}^2 = 1. \tag{4.38}$$

Из (4.37) и (4.38) следует, что

$$\boldsymbol{e}_{j1} = \sqrt{\frac{m_2}{m_1 + m_2}}\,\boldsymbol{i}_j, \quad \boldsymbol{e}_{j2} = -\sqrt{\frac{m_1}{m_1 + m_2}}\,\boldsymbol{i}_j, \tag{4.39}$$

где $\boldsymbol{i}_j$ — единичный вектор, параллельный $\boldsymbol{e}_{j1}$.

Из (4.35), (4.36) и (4.39) следует, что

$$\boldsymbol{P} = \sqrt{\frac{\omega_l^2}{4\pi\varepsilon^*}} \sum_{q,j} \boldsymbol{i}_j \left(a_{jq} e^{i\boldsymbol{q}\boldsymbol{r}} + a_{jq}^* e^{-i\boldsymbol{q}\boldsymbol{r}}\right), \tag{4.40}$$

где мы в члене $k = 2$ заменили суммирование по $\boldsymbol{q}$ на суммирование по $-\boldsymbol{q}$ и использовали (III.6.5).

Уравнение Пуассона для скалярного потенциала Φ, связанного с колебаниями кристалла, имеет вид

$$\nabla^2\Phi = -4\pi\rho = 4\pi\, d\dot{\boldsymbol{w}}\boldsymbol{P} =$$

$$= i\sqrt{\frac{4\pi\omega_l^2}{\varepsilon^*}} \sum_{q,j} (\boldsymbol{i}_j,\ \boldsymbol{q}) \left(a_{jq} e^{i\boldsymbol{q}\boldsymbol{r}} - a_{jq}^* e^{-i\boldsymbol{q}\boldsymbol{r}}\right), \tag{4.41}$$

где связанный заряд $\rho = -\operatorname{div}\boldsymbol{P}$[1]). Из этого выражения мы видим, что электрическое поле создается только продольными ко-

[1]) Тамм И. Е., гл. 2.

лебаниями, для которых $\boldsymbol{l}_j \| \boldsymbol{q}$ и, следовательно, $(\boldsymbol{l}_j, \boldsymbol{q}) = q$ (в дальнейшем мы для продольных колебаний опустим индекс j).

Уравнение (4.41) имеет решение

$$\Phi = -i\sqrt{\frac{4\pi\omega_l^2}{\varepsilon^*}}\sum_{\boldsymbol{q}}\frac{1}{q}\left(a_q e^{iqr} - a_q^* e^{-iqr}\right), \tag{4.42}$$

в чем можно убедиться, непосредственно подставив (4.42) в (4.41). Так как выражение в круглых скобках под знаком суммы чисто мнимое, то, как и должно быть, потенциал Φ вещественен.

Для определения вероятности перехода электрона, связанного с поглощением или испусканием фонона, надо вычислить соответствующий матричный элемент перехода от энергии возмущения $(-e\Phi)$. Для этого достаточно описывать состояние электрона плоской волной

$$\psi_k(\boldsymbol{r}) = \frac{1}{\sqrt{V}} e^{i\boldsymbol{k}\boldsymbol{r}}, \tag{4.43}$$

где $\boldsymbol{k}$ — волновой вектор электрона, а $V = 1\ см^3$ — объем основной области.

В матричном элементе $\langle N_q', \boldsymbol{k}' | -e\Phi | N_q, \boldsymbol{k}\rangle$ можно провести отдельно интегрирование по координатам электрона, тогда для первой экспоненты в круглой скобке в (4.42) получим

$$\frac{1}{V}\int_V e^{i(\boldsymbol{k}+\boldsymbol{q}-\boldsymbol{k}')\boldsymbol{r}}\, d\boldsymbol{r} = \delta_{\boldsymbol{k}', \boldsymbol{k}+\boldsymbol{q}},$$

что выражает закон сохранения волнового вектора электрона при поглощении фонона:

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{q}. \tag{4.44}$$

Для второй экспоненты в круглой скобке, связанной с испусканием фонона, получим

$$\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{q}. \tag{4.44а}$$

Матричные элементы от a_q и a_q^*, связанные с поглощением и испусканием фонона, согласно (III.10.25), (3.10.26) равны

$$\langle N_q' | a_q | N_q\rangle = \sqrt{\frac{\hbar N_q}{2\omega_l}}, \text{ если } N_q' = N_q - 1 \tag{4.45}$$

и

$$\langle N_q' | a_q^* | N_q\rangle = \sqrt{\frac{\hbar (N_q+1)}{2\omega_l}}, \text{ если } N_q' = N_q + 1. \tag{4.45а}$$

Таким образом, для интересующего нас матричного элемента получим

$$\langle N_q', \boldsymbol{k}' | -e\Phi | N_q, \boldsymbol{k}\rangle = \langle N_q \mp 1, \boldsymbol{k} \pm \boldsymbol{q} | -e\Phi | N_q, \boldsymbol{k}\rangle =$$
$$= \pm i\sqrt{\frac{4\pi e^2\hbar\omega_l}{\varepsilon^*}}\frac{1}{q}\cdot\begin{cases}\sqrt{N_q}, \\ \sqrt{N_q+1},\end{cases} \tag{4.46}$$

где верхние знаки (и строка) соответствуют поглощению, а нижние — испусканию фонона.

В результате взаимодействия электрона с поляризованным им кристаллом изменяется его энергия. В случае слабого взаимодействия ($\alpha \ll 1$) это изменение может быть определено посредством квантовомеханической теории возмущений. Изменение собственного значения энергии n-го квантового состояния системы, в первом и втором приближениях по энергии возмущения $\mathscr{H}'$ равно [1])

$$\mathscr{E}_n - \mathscr{E}_n^{(0)} = \langle n | \mathscr{H}' | n \rangle + \sum_{m \neq n} \frac{|\langle m | \mathscr{H}' | n \rangle|^2}{\mathscr{E}_n^{(0)} - \mathscr{E}_m^{(0)}}. \qquad (4.47)$$

Здесь $\mathscr{E}_n^{(0)}$ и $\mathscr{E}_m^{(0)}$ — невозмущенные энергии n-го и m-го состояний.

Рассмотрим изменение энергии свободного электрона $\varepsilon^{(0)} = \hbar^2 k^2/2m^*$ при взаимодействии его с поляризованным кристаллом. Если абсолютная температура кристалла равна нулю ($N_q = 0$), то электрон может только испускать фононы ($N'_q = N_q + 1 = 1$; $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{q}$). Так как при $N'_q = N_q$ матричный элемент (4.46) равен нулю ($\langle n | \mathscr{H}' | n \rangle = 0$), то энергия электрона изменяется только во втором приближении теории возмущений. Из (4.47) и (4.46) следует, что

$$\varepsilon(\boldsymbol{k}) - \varepsilon^{(0)} = \sum_{\boldsymbol{q}} \frac{|\langle 1, \boldsymbol{k}-\boldsymbol{q} | -e\Phi | 0, \boldsymbol{k} \rangle|^2}{\frac{\hbar^2 k^2}{2m^*} - \frac{\hbar^2 (\boldsymbol{k}-\boldsymbol{q})^2}{2m^*} - \hbar\omega_l} =$$

$$= -\frac{2\pi e^2 \hbar \omega_l}{\varepsilon^*} \int \frac{d\tau_q}{(2\pi)^3} \frac{1}{q^2} \frac{1}{\hbar\omega_l - \frac{\hbar^2}{2m^*}(2\boldsymbol{k}\boldsymbol{q} - \boldsymbol{q}^2)}, \qquad (4.48)$$

где мы перешли от суммирования по $\boldsymbol{q}$ к интегрированию. Промежуточному состоянию m соответствует фотон $\hbar\omega_l$ и волновой вектор электрона $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{q}$; напомним, что при переходе из основного состояния n в промежуточное m закон сохранения энергии, вообще говоря, не выполняется.

Введем в интеграле полярные координаты с осью, направленной вдоль $\boldsymbol{k}$ ($d\tau_q = 2\pi q^2 dq \sin\vartheta\, d\vartheta$; $\boldsymbol{k}\boldsymbol{q} = kq\cos\vartheta$), и разложим подынтегральное выражение в ряд по степеням k, до k^2 включительно. Интегрирование по ϑ первого слагаемого (не зависящего от k) дает множитель 2, второго слагаемого (пропорционального k) — множитель нуль и третьего слагаемого (пропорционального k^2) — множитель $^2/_3$. В результате интегралы по q от 0 до ∞ от первого и третьего слагаемых сводятся к интегралам

---

[1]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика, § 38.

от рациональных дробей; они вычисляются элементарно или могут быть взяты из таблиц [1]).

В результате из (4.48) мы получим

$$\varepsilon(k)=\frac{\hbar^2k^2}{2m^*}-\alpha\left(\hbar\omega_l+\frac{\hbar k^2}{12m^*}\right)=-\alpha\hbar\omega_l+\frac{\hbar^2k^2}{2m_{\text{pol}}}, \qquad (4.49)$$

где эффективная масса полярона

$$m_{\text{pol}}=\frac{m^*}{1-\frac{\alpha}{6}}\approx m^*\left(1+\frac{\alpha}{6}\right). \qquad (4.49\text{а})$$

Здесь α — параметр, введенный в (4.31).

Мы видим, что для полярона слабой связи уровень энергии основного состояния понижается на величину $\alpha\hbar\omega_l$, а его эффективная масса m_{pol} увеличивается в отношении $\frac{m_{\text{pol}}}{m^*}\approx 1+\frac{\alpha}{6}$.

Типичные значения α для полупроводников с частично ионной связью меньше единицы. Для соединений $A_{III}B_V$ α лежит в пределах от 0,015 (InSb) до 0,080 (InP), для соединений $A_{II}B_{VI}$ — от 0,39 (CdTe) до 0,65 (CdS). При вычислении α для этих соединений были использованы малые эффективные массы электронов m^*. Для щелочно-галоидных кристаллов $\alpha>1$; например, для LiI $\alpha=2{,}4$, для RbBr $\alpha=6{,}6$, при вычислениях для электрона была использована масса свободного электрона ($m=0{,}9\cdot10^{-27}$ *г*). Поскольку мы пользовались как в случае сильной, так и слабой связи континуальным приближением, изложенные выше соображения относятся только к поляронам большого радиуса, когда радиус полярона больше (точнее много больше) постоянной решетки.

Мы не рассматриваем теорию поляронов малого радиуса, когда радиус полярона равен или меньше постоянной решетки [2]).

---

[1]) Бронштейн П. Н., Семендяев К. А. Справочник по математике, — 11 изд. — М., Наука, 1967, с. 351.

[2]) Читатель может обратиться к сборнику Поляроны. Под ред. Фирсова Ю. А. — М., 1975, в котором статья Ю. А. Фирсова специально посвящена поляронам малого радиуса.

ГЛАВА VI

ЭЛЕКТРИЧЕСКИЕ, ТЕПЛОВЫЕ И МАГНИТНЫЕ СВОЙСТВА ТВЕРДЫХ ТЕЛ

§ 1. Металлы, диэлектрики и полупроводники

В гл. IV мы рассмотрели поведение отдельного электрона в периодическом поле идеального кристалла. Было показано, что электрону в стационарном состоянии соответствует незатухающая средняя скорость (IV.3.28), т. е. он свободно перемещается по всему объему кристалла. На первый взгляд все вещества должны были бы быть очень хорошими проводниками (металлами) с числом электронов проводимости, равным полному числу электронов в теле. На самом деле, даже в металлах число электронов проводимости значительно меньше — порядка числа атомов, а в диэлектриках — равно нулю (при абсолютном нуле температуры).

Для того чтобы понять эту ситуацию, рассмотрим совокупность всех NsZ электронов твердого тела, где N — число кристаллических ячеек, s — число атомов в кристаллической ячейке и Z — число электронов атома (атомный номер) [1]). Взаимодействие между электронами мы учитываем только в среднем посредством самосогласованного поля, но мы учтем принцип Паули, согласно которому в одном квантовом состоянии (характеризующемся волновым вектором $\boldsymbol{k}$) не может находиться больше двух электронов и только с противоположно направленными спинами. При абсолютном нуле температуры, когда система находится в наинизшем энергетическом состоянии, электроны твердого тела должны занять наинизшие $NsZ/2$ квантовых состояний, подробно рассмотренных в гл. IV. При этом возможны два случая: либо наивысший, заполненный электронами уровень энергии совпадет с верхним краем одной из разрешенных зон энергии (рис. VI.1, *а*), либо он попадет внутрь такой зоны (рис. VI.1, б). В последнем случае, если приложить к телу даже слабое электрическое поле, то электроны, расположенные вблизи границы ε_0, будут уско-

[1]) Для простоты мы рассматриваем простое (одноатомное) вещество.

ряться и переходить в другие более высокие квантовые состояния, непрерывно примыкающие к ε_0, не занятые другими электронами. В результате число электронов, движущихся по и против поля, не будет одинаково — возникнет электрический ток; тело будет вести себя как металл. В первом случае, когда электроны полностью заполняют верхнюю энергетическую зону, электрическое поле не может вызвать перераспределения электронов, по крайней мере до тех пор, пока оно не слишком велико (меньше 10^6 *в/см*) и не вызывает переброса электронов в более высокие разрешенные зоны энергии (пробой). В этом случае электрический ток возникнуть не может, тело ведет себя как диэлектрик (изолятор). Также ведут себя и электроны более глубоких, полностью заполненных энергетических зон в металле. Такое простое качественное объяснение отличия между металлом и диэлектриком явилось большим триумфом зонной теории.

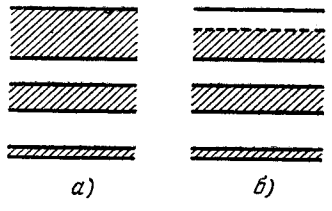


Рис. VI.1.

Каждая отдельная зона энергии содержит $2N$ квантовых состояний; поэтому для простого кристалла (с $s=1$) число заполненных электронами зон равно $NZ:2N=Z/2$. Таким образом, если Z нечетно, то заполнено нецелое число зон, т. е. элементы с нечетным атомным номером, кристаллизующиеся в простой решетке, — металлы. Обратное заключение, что элементы с четным Z (при $s=1$) ведут себя как диэлектрики, неправильно, так как всегда возможно перекрытие энергетических зон (гл. IV, § 5), а в этом случае тело будет вести себя как металл. Более сложная структура решетки (меньшая симметрия) и в особенности наличие атомов разных сортов способствуют разделению энергетических зон, т. е. образованию диэлектрика. Твердый водород — диэлектрик, хотя на первый взгляд он должен был бы вести себя как щелочный металл. Такое поведение водорода и ряда других элементов связано с тем, что при кристаллизации образуются молекулы вещества, которые связаны в решетке ван-дер-ваальсовскими силами ($s\neq 1$). С зонной точки зрения мы должны, конечно, в этом случае иметь полностью заполненную электронами полосу энергии. Таким образом, при абсолютном нуле температуры все твердые тела ведут себя либо как металлы, либо как диэлектрики.

В случае диэлектрика, когда электроны полностью заполняют самую верхнюю зону (валентную), отделенную от следующей полосы разрешенных энергий (зоны проводимости) запрещенной зоной шириною ε_g, электроны будут при повышении температуры переходить из валентной зоны в зону проводимости.

Если ε_G порядка 1 *эв* или меньше, то при комнатных температурах и выше чистые кристаллы вещества обнаруживают заметную электропроводность, обусловленную движением электронов в зоне проводимости и дырок в валентной зоне.

Такие чистые вещества с узкой запретной зоной получили название *собственных полупроводников*. Практически чаще приходится иметь дело с *примесными полупроводниками*, когда поставщиками электронов в зоне проводимости являются доноры, а поставщиками дырок в валентной зоне — акцепторы гл. V, § 2, п. 1.

Добавляя различные примесные атомы в разных количествах к чистому веществу, можно получить полупроводники с весьма разнообразными электрическими свойствами. Такие процессы легирования или допирования полупроводника различными примесями наиболее широко были исследованы для германия и кремния, которые получили в последние годы большое техническое применение.

§ 2. Статистическое равновесие свободных электронов в полупроводниках и металлах

1. В настоящей главе мы рассмотрим некоторые свойства полупроводников и металлов, обусловленные свободными электронами (дырками), находящимися в состоянии статистического (термодинамического) равновесия с тепловыми колебаниями кристаллической решетки. Важной особенностью систем, находящихся в состоянии статистического (термодинамического) равновесия, является то, что их свойства не зависят от механизма взаимодействия, приводящего к этому равновесию. Поэтому в нашем случае нет необходимости рассматривать конкретный механизм взаимодействия свободных электронов и дырок с колебаниями решетки и процессы теплового возбуждения и рекомбинации электронов (дырок). Как мы увидим в следующих главах, эти механизмы существенны при рассмотрении кинетических явлений, т. е. процессов электро- и теплопроводимости, гальвано- и термомагнитных явлений и т. д.

В одноэлектронном приближении взаимодействие между электронами кристалла учитывается только посредством самосогласованного поля, в котором каждый электрон движется независимо от других. Со статистической точки зрения одноэлектронное приближение соответствует модели идеального газа. В состоянии статистического равновесия идеальный газ электронов подчиняется статистике Ферми—Дирака. В статистическом равновесии среднее число электронов в определенном квантовом состоянии, характеризующемся тремя квантовыми числами

k_1, k_2, k_3 [1]), с энергией ε_k при температуре T равно [2])

$$f_0(\varepsilon_k) = \frac{1}{\exp\frac{\varepsilon_k - \zeta}{k_0 T} + 1}. \tag{2.1}$$

Здесь ζ — химический потенциал в расчете на 1 электрон, k_0 — постоянная Больцмана. Функция $f_0(\varepsilon_k)$ называется *функцией распределения Ферми — Дирака*.

Полное число электронов в объеме V равно

$$N = \sum_k f_0(\varepsilon_k) = \sum_k \frac{1}{\exp\frac{\varepsilon_k - \zeta}{k_0 T} + 1}, \tag{2.2}$$

где суммирование ведется по всем квантовым состояниям электрона в объеме V с учетом спина, т. е. с учетом того, что в каждом о р б и т а л ь н о м квантовом состоянии могут находиться (в соответствии с принципом Паули) два электрона с противоположно направленными спинами.

Равенство (2.2) определяет химический потенциал ζ как функцию концентрации электронов $n = N/V$ и температуры T. Если число квантовых состояний электрона (без учета спина) в 1 $см^3$ в интервале энергии $d\varepsilon$ равно $g(\varepsilon)\,d\varepsilon$, то число электронов в единице объема с энергией между ε и $\varepsilon + d\varepsilon$ равно

$$n(\varepsilon)\,d\varepsilon = 2 f_0(\varepsilon)\, g(\varepsilon)\, d\varepsilon. \tag{2.3}$$

Если состояние электронов в кристалле характеризуется волновым вектором $\boldsymbol{k}$ ($k_1 = k_x$, $k_2 = k_y$, $k_3 = k_z$), то в некоторых случаях удобнее пользоваться не функцией $g(\varepsilon)$, а числом квантовых состояний электрона, конец волнового вектора которого $\boldsymbol{k}$ лежит в элементе $d\tau_k$ $\boldsymbol{k}$-пространства. Это число квантовых состояний в 1 $см^3$, согласно (III.5.27), равно $d\tau_k/(2\pi)^3$ [3]). Число электронов, у которых составляющие волнового вектора лежат в интервалах от k_x до $k_x + dk_x$, k_y до $k_y + dk_y$, k_z до $k_z + dk_z$, равно

$$n(\boldsymbol{k})\,dk_x\,dk_y\,dk_z = 2 f_0(\boldsymbol{k})\frac{dk_x\,dk_y\,dk_z}{(2\pi)^3}. \tag{2.3а}$$

Здесь $f_0(\boldsymbol{k})$ — функция распределения (2.1), в которой энергия

---

[1]) k_1, k_2, k_3 — три дискретных или квазидискретных квантовых числа, характеризующих о р б и т а л ь н о е движение электрона; очень часто мы будем под ними понимать три составляющих волнового вектора электрона $\boldsymbol{k}$ ($k_1 = k_x$, $k_2 = k_y$, $k_3 = k_z$); для электрона в атоме им будут соответствовать: главное квантовое число n, азимутальное квантовое число l и магнитное квантовое число m.

[2]) А н с е л ь м А. И., гл. IX, § 2.

[3]) В пространстве квазиимпульса $\boldsymbol{p} = \hbar\boldsymbol{k}$, аналогичное число квантовых состояний отличается множителем $\hbar^{-3}$.

$\varepsilon_{\boldsymbol{k}} = \varepsilon(\boldsymbol{k})$, выражена через волновой вектор $\boldsymbol{k}$ (в приближении скалярной эффективной массы $\varepsilon(\boldsymbol{k}) = \hbar^2 k^2/2m^* = p^2/2m^*$).

Мы видели, что для электронов в периодическом поле $g(\varepsilon)$ определяется выражением (IV.3.25). Если энергия электрона $\varepsilon = \frac{1}{2m^*} p^2$, то

$$g(\varepsilon) = \frac{\sqrt{2}}{2\pi^2} \frac{m^{*3/2}}{\hbar^3} \sqrt{\varepsilon}\,. \tag{2.4}$$

Таким образом, в этом случае концентрация электронов

$$n = 2\int_0^\infty f_0(\varepsilon)\, g(\varepsilon)\, d\varepsilon = \frac{\sqrt{2}}{\pi^2} \frac{m^{*3/2}}{\hbar^3} \int_0^\infty \frac{\varepsilon^{1/2}\, d\varepsilon}{\exp \frac{\varepsilon - \zeta}{k_0 T} + 1} =$$

$$= \frac{\sqrt{2}}{\pi^2} \frac{(m^* k_0 T)^{3/2}}{\hbar^3} \int_0^\infty \frac{x^{1/2}\, dx}{e^{x-z}+1}\,, \tag{2.5}$$

где $x = \varepsilon/k_0 T$ и $z = \zeta/k_0 T$. Это выражение определяет ζ как функцию n, T и m^*.

Интеграл, стоящий в правой части (2.5), не выражается через элементарные функции от z. Интегралы вида

$$\mathscr{F}_n(z) = \int_0^\infty \frac{x^n\, dx}{e^{x-z}+1} \tag{2.6}$$

часто встречаются в электронной теории кристаллов (в уравнении (2.5) $n = {}^1/_2$). Для случая $n = -{}^1/_2, {}^1/_2, {}^3/_2$ интегралы (2.6) были табулированы Мак-Дугаллом и Стонером [1]; для $n = {}^5/_2, {}^7/_2, {}^9/_2, {}^{11}/_2$ — Биром с сотрудниками [2]) и для $n = 1, 2, 3, 4$ — Родсом [3]).

Рассмотрим два важных предельных случая, когда интеграл в выражении (2.5) может быть просто вычислен.

2. Рассмотрим случай сильно вырожденного электронного газа, когда $z \gg 1$, т. е. химический потенциал $\zeta \gg k_0 T$. Как мы увидим ниже, этот случай реализуется в хороших металлах, у которых концентрация электронов проводимости $n \approx \approx 10^{22}$ *см*$^{-3}$. Очевидно, что в этом случае $f_0(\varepsilon) \approx 1$ для $\varepsilon \ll \zeta$ и $f_0(\varepsilon) \approx 0$ для $\varepsilon \gg \zeta$; крутой спад $f_0(\varepsilon)$ происходит вблизи точки $\varepsilon = \zeta$ в интервале шириной порядка $k_0 T$.

Функции $f_0(\varepsilon)$ и $\left(-\frac{\partial f_0}{\partial \varepsilon}\right)$ схематически изображены на рис. VI.2. Чем ниже температура, тем $\left(-\frac{\partial f_0}{\partial \varepsilon}\right)$ ближе к дельта-

[1]) Mc. Dougall T., Stoner E. C.—Phil. Trans. Roy. Soc., 1938, v. A 237, p. 350.

[2]) Beer A. C., Chase M. N., Choquard P. F.—Helv. Phys. Acta, 1955, v. 28, p. 529.

[3]) Rhodes P.—Proc. Roy. Soc., 1950, v. A 204, p. 396.

функции. В самом деле, $\left(-\frac{\partial f_0}{\partial \varepsilon}\right)$ отлична от нуля только вблизи точки $\varepsilon=\zeta$ и

$$\int_0^\infty \left(-\frac{\partial f_0}{\partial \varepsilon}\right) d\varepsilon = -\int_0^\infty df_0 = f_0(0) - f_0(\infty) = 1,$$

так как $f_0(0)=1$ и $f_0(\infty)=0$.

Таким образом, $\left(-\frac{\partial f_0}{\partial \varepsilon}\right)=\delta(\varepsilon-\zeta_0)$, где ζ_0 — значение ζ при $T=0$.

Вычислим в случае сильного вырождения интеграл ($\chi(\varepsilon)$ — произвольная функция):

$$I=\int_0^\infty \chi(\varepsilon) f_0(\varepsilon)\, d\varepsilon = \int_0^\infty f_0(\varepsilon)\, d\varphi(\varepsilon) =$$

$$= f_0(\varepsilon)\varphi(\varepsilon)\Big|_0^\infty - \int_0^\infty \varphi(\varepsilon)\frac{\partial f_0}{\partial \varepsilon}\, d\varepsilon = -\varphi(0) + \int_0^\infty \varphi(\varepsilon)\left(-\frac{\partial f_0}{\partial \varepsilon}\right) d\varepsilon, \quad (2.7)$$

где на первом этапе выполнено преобразование $\chi(\varepsilon)\,d\varepsilon = d\varphi(\varepsilon)$, обычное при интегрировании по частям. Как мы увидим ниже, $\varphi(0)=0$ для всех практических случаев, поэтому

$$I=\int_0^\infty \varphi(\varepsilon)\left(-\frac{\partial f_0}{\partial \varepsilon}\right) d\varepsilon. \quad (2.7а)$$

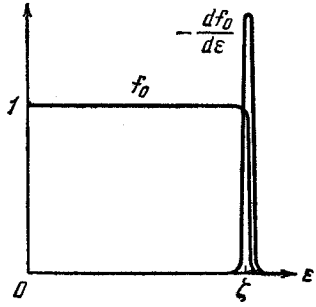


Рис. VI.2.

Используем эту формулу для вычисления выражения (2.5) в приближении, когда $\left(-\frac{\partial f_0}{\partial \varepsilon}\right)\approx\delta(\varepsilon-\zeta_0)$.

Так как в (2.5) $\chi(\varepsilon)=2g(\varepsilon)$, то $\varphi(\varepsilon)=\frac{\sqrt{2}}{\pi^2}\frac{m^{*3/2}}{\hbar^3}\frac{2}{3}\varepsilon^{3/2}$. Таким образом,

$$n=\frac{\sqrt{2}}{\pi^2}\frac{m^{*3/2}}{\hbar^3}\frac{2}{3}\int_0^\infty \varepsilon^{3/2}\delta(\varepsilon-\zeta_0)\, d\varepsilon = \frac{(2m^*)^{3/2}}{2\pi^2\hbar^3}\zeta_0^{3/2}, \quad (2.8)$$

откуда химический потенциал

$$\zeta_0=\frac{\hbar^2}{2m^*}(3\pi^2 n)^{2/3}=\frac{(2\pi\hbar)^2}{2m^*}\left(\frac{3n}{8\pi}\right)^{2/3}, \quad (2.8а)$$

что в точности совпадает с максимальной кинетической энергией электронов ε_0 идеального ферми-газа при абсолютном нуле температуры (Приложение 4).

В следующем приближении (Приложение 17) химический потенциал

$$\zeta = \zeta_0 \left[1 - \frac{\pi^2}{12} \left(\frac{k_0 T}{\zeta_0} \right)^2 \right]. \qquad (2.9)$$

Используя нулевое приближение для химического потенциала, запишем условие сильного вырождения в явном виде:

$$\frac{\zeta_0}{k_0 T} = \frac{\hbar^2 (3\pi^2 n)^{2/3}}{2m^* k_0 T} \gg 1. \qquad (2.10)$$

Как видно из функции распределения $f_0(\varepsilon)$, условие вырождения (2.10) носит на самом деле экспоненциальный характер, т. е. $\exp(\zeta_0/k_0T) \gg 1$; поэтому, если $\zeta_0/k_0T \approx 5 \div 7$, то вырождение можно считать сильным.

Из (2.10) мы видим, что вырождению способствует высокая концентрация n, малая эффективная масса m^* и низкая температура T. Для типичного металла с $n \approx 10^{22}$ *см*$^{-3}$, $m^* \approx 10^{-27}$ *г* при комнатной температуре $\zeta_0/k_0T \approx 10^2$, т. е. вырождение очень сильное. Легко видеть, что свободные электроны в металле остаются сильно вырожденными вплоть до его температуры плавления.

3. Мы рассмотрели случай положительного химического потенциала ζ, удовлетворяющего неравенству $\exp(\zeta/k_0T) \gg 1$ (условие вырождения). Рассмотрим теперь противоположный случай — отрицательного химического потенциала, удовлетворяющего неравенству $\exp(-\zeta/k_0T) \gg 1$. В этом случае функция распределения

$$f_0(\varepsilon) = \frac{1}{e^{(\varepsilon-\zeta)/k_0T} + 1} \approx e^{\zeta/k_0T}\, e^{-\varepsilon/k_0T} = Ae^{-\varepsilon/k_0T}, \qquad (2.11)$$

т. е. переходит в распределение Максвелла — Больцмана с нормировочной константой $A \equiv \exp(\zeta/k_0T)$. Концентрация электронов в зоне проводимости, согласно (2.5), равна

$$n = \frac{\sqrt{2}}{\pi^2} \frac{m^{*3/2}}{\hbar^3} \int_0^\infty Ae^{-\varepsilon/k_0T} \varepsilon^{1/2} d\varepsilon = \frac{(2\pi m^* k_0 T)^{3/2}}{4\pi^3 \hbar^3} A. \qquad (2.12)$$

При вычислении интеграла мы ввели переменную $x = \varepsilon^{1/2}$ и воспользовались формулой (П.7.2). Из формулы (2.12) следует

$$A = e^{\zeta/k_0T} = \frac{4\pi^3 \hbar^3 n}{(2\pi m^* k_0 T)^{3/2}}, \qquad (2.12а)$$

$$\zeta = k_0 T \ln \left[\frac{4\pi^3 \hbar^3 n}{(2\pi m^* k_0 T)^{3/2}} \right]. \qquad (2.12б)$$

В противоположность вырожденному случаю (2.9), в классической статистике химический потенциал ζ, как видно из (2.12б), довольно сильно зависит от температуры.

Функция распределения (2.11) может быть записана теперь в следующем виде:

$$f_0(\varepsilon) = \frac{4\pi^3\hbar^3 n}{(2\pi m^* k_0 T)^{3/2}} e^{-\varepsilon/k_0 T}. \qquad (2.11а)$$

Критерий применимости классической статистики имеет вид

$$\exp\left(-\frac{\zeta}{k_0 T}\right) = \frac{1}{A} = \frac{(2\pi m^* k_0 T)^{3/2}}{4\pi^3\hbar^3 n} \gg 1, \qquad (2.13)$$

что согласуется с (2.10).

Мы видим, что применению классической статистики способствует малая концентрация n, высокая температура T и большая эффективная масса m^*. Для $n \approx 10^{17}$ *см*$^{-3}$, $m^* \approx 10^{-27}$ *г* при комнатной температуре $\frac{1}{A} \approx 300$, т. е. критерий (2.13) выполняется с избытком. Граничное значение концентрации, соответствующее $\frac{1}{A} \approx 1$, равно $n \approx 10^{19}$ *см*$^{-3}$.

4. Обычно в полупроводниках концентрация электронов в зоне проводимости сама является функцией температуры. Это происходит в силу теплового возбуждения электронов примесей и валентной зоны. Формулы, полученные выше, остаются, конечно, правильными, но мало эффективными, так как нам неизвестна явная зависимость $n(T)$. Кроме того, при возбуждении электронов из валентной зоны в ней остаются положительно заряженные подвижные дырки, которые наряду с электронами зоны проводимости участвуют в явлениях переноса. Поэтому в этом случае задача должна быть рассмотрена иначе. На рис. VI.3 представлена схема уровней энергии примесного полупроводника с запрещенной зоной шириной ε_G и донорными и акцепторными уровнями, отстоящими от нижнего края зоны проводимости на величину ε_D и ε_A. Все энергии условимся отсчитывать от нулевого уровня, совпадающего с нижним краем зоны проводимости. Вероятность того, что квантовое состояние с энергией ε не занято электроном, т. е. по определению является дыркой, равна

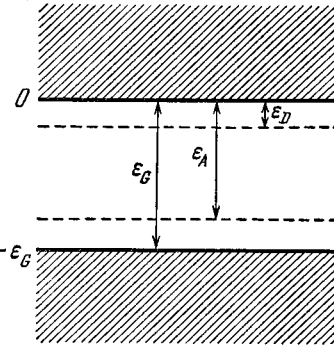


Рис. VI. 3.

$$f_0'(\varepsilon) = 1 - f_0(\varepsilon) = 1 - \frac{1}{e^{\frac{\varepsilon-\zeta}{k_0 T}} + 1} = \frac{1}{e^{\frac{\zeta-\varepsilon}{k_0 T}} + 1}, \qquad (2.14)$$

где $f_0'(\varepsilon)$ — функция распределения Ферми — Дирака для дырок.

Если отсчитывать все энергии от нижнего края зоны проводимости, то энергия электрона в зоне проводимости равна $\varepsilon = \hbar^2 k^2/2m_n$, на донорном уровне $\varepsilon = -\varepsilon_D$, на акцепторном уровне $\varepsilon = -\varepsilon_A$ и в валентной зоне $\varepsilon = -\varepsilon_G - \varepsilon'$, где $\varepsilon' = \hbar^2 k'^2/2m_p$ — «кинетическая» энергия дырки, а ε_G — ширина запрещенной зоны.

Если ввести «химический потенциал дырок» $\zeta' = -\varepsilon_G - \zeta$, то функция распределения для дырок (2.14) может быть записана в виде

$$f_0'(\varepsilon) = \frac{1}{\exp\frac{\varepsilon' - \zeta'}{k_0 T} + 1} \qquad (2.14а)$$

— совершенно аналогично функции распределения (2.1) для электронов.

Условие нейтральности полупроводника, из которого определяется его химический потенциал ζ, может быть выражено в следующей форме: (число электронов в зоне проводимости) + + (число электронов на акцепторных уровнях) = (число дырок в валентной зоне) + (число дырок на донорных уровнях), т. е.

$$\int\limits_{\text{(зона пров.)}} \frac{2g(\varepsilon)\,d\varepsilon}{\exp\frac{\varepsilon - \zeta}{k_0 T} + 1} + \sum_{(A)} \frac{1}{\exp\frac{-\varepsilon_A - \zeta}{k_0 T} + 1} = \int\limits_{\text{(вал. зона)}} \frac{2g(\varepsilon')\,d\varepsilon'}{\exp\frac{\zeta + \varepsilon_G + \varepsilon'}{k_0 T} + 1} + \sum_{(D)} \frac{1}{\exp\frac{\zeta + \varepsilon_D}{k_0 T} + 1}. \qquad (2.15)$$

Если электроны в зоне проводимости и дырки в валентной зоне подчиняются классической статистике: $A \equiv \exp(\zeta/k_0 T) \ll 1$ и $\exp\frac{\zeta + \varepsilon_G}{k_0 T} \gg 1$, то интегралы в (2.15) аналогичны интегралу в (2.12), так что

$$\frac{(2\pi m_n k_0 T)^{3/2}}{4\pi^3 \hbar^3} A + \frac{n_A}{\frac{1}{A}\exp\left(-\frac{\varepsilon_A}{k_0 T}\right) + 1} = \frac{(2\pi m_p k_0 T)^{3/2}}{4\pi^3 \hbar^3} \frac{1}{A} e^{\frac{-\varepsilon_G}{k_0 T}} + \frac{n_D}{A\exp\left(+\frac{\varepsilon_D}{k_0 T}\right) + 1}. \qquad (2.15а)$$

Здесь n_A и n_D — концентрации акцепторов и доноров. Из (2.15а) видно, что концентрации электронов проводимости n и дырок в валентной зоне p равны

$$n = \frac{(2\pi m_n k_0 T)^{3/2}}{4\pi^3 \hbar^3} \exp\frac{\zeta}{k_0 T}, \qquad (2.15б)$$

$$p = \frac{(2\pi m_p k_0 T)^{3/2}}{4\pi^3 \hbar^3} \exp\left(-\frac{\varepsilon_G + \zeta}{k_0 T}\right). \qquad (2.15в)$$

Уравнение нейтральности (2.15а) определяет величину A или химический потенциал $\zeta = k_0 T \ln A$. В общем случае это алгеб-

раическое уравнение 4-й степени относительно A. Существуют эффективные численные и графические методы его решения. Мы ограничимся рассмотрением ряда частных, но важных случаев, когда решение может быть получено в виде простых формул.

1) Собственный полупроводник ($n_A = n_D = 0$).

Из (2.15а) получим

$$A = \left(\frac{m_p}{m_n}\right)^{3/4} \exp\left(-\frac{\varepsilon_G}{2k_0T}\right), \tag{2.16}$$

откуда

$$\zeta = -\frac{\varepsilon_G}{2} + \frac{3}{4} k_0T \ln\left(\frac{m_p}{m_n}\right). \tag{2.16а}$$

Так как множитель $\frac{3}{4}\ln\left(\frac{m_p}{m_n}\right) \sim 1$, то второе слагаемое в (2.16а) порядка k_0T и, следовательно, с такой точностью химический потенциал ζ в собственном полупроводнике (диэлектрике) находится в середине запрещенной зоны. Если $m_p = m_n$, то это утверждение точно, и в этом случае ζ от температуры не зависит (в той мере, в какой мы не рассматриваем зависимость ширины зоны ε_G от температуры). Концентрации электронов в зоне проводимости и дырок в валентной зоне одинаковы и равны

$$n_i = \frac{(2\pi\sqrt{m_n m_p}k_0T)^{3/2}}{4\pi^3\hbar^3} \exp\left(-\frac{\varepsilon_G}{2k_0T}\right) = n_T^0 \exp\left(-\frac{\varepsilon_G}{2k_0T}\right). \tag{2.16б}$$

Легко видеть, что предъэкспоненциальный множитель n_T^0 по порядку величины равен концентрации, соответствующей одной частице на объем λ^3, где λ—длина дебройлевской волны электрона (дырки), движущегося с тепловой скоростью. Для $m_n = m_p = m = 0{,}9 \cdot 10^{-27}$ г и $T = 294°$ К $n_T^0 = 2{,}44 \cdot 10^{19}$ см$^{-3}$. Концентрация n_T^0 неограниченно возрастает при $T \to \infty$ и не ограничена числом электронов в валентной зоне. Это связано с тем, что при выводе уравнения (2.15а) нигде не учитывается конечная ширина валентной зоны. Условие нормировки в случае собственного полупроводника заключается только в утверждении, что число электронов в зоне проводимости равно числу свободных мест (дырок) в валентной зоне.

2) Примесный полупроводник (например, донорного типа $n_A = 0$) с широкой запрещенной зоной ($\varepsilon_G \gg \varepsilon_D$).

Из (2.15а) имеем

$$\frac{(2\pi m_n k_0T)^{3/2}}{4\pi^3\hbar^3} A\left(Ae^{\frac{\varepsilon_D}{k_0T}} + 1\right) = n_D, \tag{2.17}$$

где мы пренебрегли членом, пропорциональным $\exp\left(-\frac{\varepsilon_G}{k_0T}\right)$ ($\varepsilon_G \gg \varepsilon_D$!).

Рассмотрим два случая.

а) $Ae^{\frac{\varepsilon_D}{k_0T}} \gg 1$.

Так как $A \ll 1$, то это может иметь место только в случае $\varepsilon_D \gg k_0T$. Пренебрегая в скобке выражения (2.17) единицей, получим

$$A = \frac{(2\pi\hbar)^{3/2} n_D^{1/2}}{\sqrt{2}\,(2\pi m_n k_0 T)^{3/4}} e^{-\frac{\varepsilon_D}{2k_0T}} \tag{2.18}$$

и

$$\zeta = -\frac{\varepsilon_D}{2} + \frac{1}{2} k_0 T \ln \frac{(2\pi\hbar)^3 n_D}{2\,(2\pi m_n k_0 T)^{3/2}}. \tag{2.18а}$$

При концентрации доноров $n_D = 10^{17}$ *см*$^{-3}$, $m_n = 10^{-27}$ *г* и температуре $T = 300°$ К величина, стоящая под логарифмом в (2.18а) порядка 10^{-2}, т. е. второе слагаемое порядка k_0T. Это означает, что химический потенциал проходит примерно посредине между нижним краем зоны проводимости и уровнем донора.

Концентрация электронов проводимости

$$n = \frac{2^{1/2}\,(2\pi m_n k_0 T)^{3/4}\, n_D^{1/2}}{(2\pi\hbar)^{3/2}} e^{-\frac{\varepsilon_D}{2k_0T}}. \tag{2.18б}$$

Заметим, что $n \sim \sqrt{n_D}$.

б) $Ae^{\frac{\varepsilon_D}{k_0T}} \ll 1$.

Так как $A \ll 1$, то этот случай всегда реализуется, если $\varepsilon_D \leqslant k_0T$. Пренебрежем в этом случае в скобке (2.17) величиной $A \exp \frac{\varepsilon_D}{k_0T}$; получим

$$A = \frac{4\pi^3\hbar^3 n_D}{(2\pi m_n k_0 T)^{3/2}}, \tag{2.19}$$

$$\zeta = k_0 T \ln \frac{4\pi^3\hbar^3 n_D}{(2\pi m_n k_0 T)^{3/2}}. \tag{2.19а}$$

Концентрация электронов проводимости

$$n = n_D. \tag{2.19б}$$

Этот результат имеет вполне наглядный смысл: если $k_0T \gtrsim \varepsilon_D$, то практически все доноры ионизованы и $n \approx n_D$. Легко видеть, что химический потенциал ζ при этом отрицателен и расположен ниже донорного уровня, т. е. $|\zeta| > \varepsilon_D$.

Если $\varepsilon_D \gg k_0T$, то можно всегда подобрать такие малые n_D, чтобы выполнялось условие б). Например, если $\varepsilon_D = 10k_0T$, то для того, чтобы все доноры были ионизованы при комнатной температуре, необходимо, чтобы $n_D \ll 10^{15}$ *см*$^{-3}$. Это тоже имеет вполне наглядный смысл, так как малая концентрация доноров весьма затрудняет переходы на них электронов из зоны проводимости.

Аналогично случаю 2) может быть рассмотрен примесный полупроводник акцепторного типа ($n_D = 0$).

Если учитывать в статистическом равновесии электронов одновременно зону проводимости, валентную зону и один сорт примесей (например, доноры), то уравнение (2.15а) превращается в алгебраическое уравнение 3-й степени относительно A и может решаться численно или графически.

Перемножая (2.15б) и (2.15в) и сравнивая с (2.16б), получим полезную формулу

$$np = \frac{(2\pi \sqrt{m_n m_p} k_0 T)^3}{(4\pi^3 \hbar^3)^2} e^{-\frac{\varepsilon_G}{k_0 T}} = n_i^2, \tag{2.20}$$

т. е. произведение концентраций электронов и дырок в примесном полупроводнике равно квадрату концентрации электронов (или дырок) в том же полупроводнике, лишенном примесей (собственном полупроводнике).

5. Для того чтобы применить уравнение (2.15) к полупроводникам со сложной структурой энергетической зоны, таким, например, как германий или кремний, надо определить в этом случае плотность состояний $g(\varepsilon)$.

В n-германии и n-кремнии энергетический спектр электронов состоит из ряда эквивалентных минимумов, симметрично расположенных в зоне Бриллюэна. Вблизи каждого из этих минимумов энергия

$$\varepsilon(\boldsymbol{k}) = \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_1} + \frac{k_y^2}{m_2} + \frac{k_z^2}{m_3}\right), \tag{2.21}$$

где $\boldsymbol{k}$ — волновой вектор электрона, отсчитываемый от точки минимума; оси x, y и z направлены по главным осям эллипсоида энергии $\varepsilon(\boldsymbol{k}) = \text{const}$ (для германия и кремния поверхности постоянной энергии — эллипсоиды вращения, так что $m_1 = m_2$).

Введем в $\boldsymbol{k}$-пространстве преобразование

$$k_x = k_x' \sqrt{m_1},\ k_y = k_y' \sqrt{m_2},\ k_z = k_z' \sqrt{m_3}. \tag{2.21а}$$

Тогда энергия

$$\varepsilon = \hbar^2 k'^2/2. \tag{2.21б}$$

Число состояний в 1 $см^3$ в элементарном объеме $dk_x\, dk_y\, dk_z$ равно

$$\frac{dk_x\, dk_y\, dk_z}{8\pi^3} = (m_1 m_2 m_3)^{1/2} \frac{dk_x' dk_y' dk_z'}{8\pi^3}.$$

Учитывая (2.21б), получим аналогично (IV.3.27)

$$g(\varepsilon) = \frac{\sqrt{2}}{2\pi^2} N_c \frac{(m_1 m_2 m_3)^{1/2}}{\hbar^3} \sqrt{\varepsilon}, \tag{2.22}$$

где N_c — число эквивалентных минимумов. Для германия $N_c = 4$, для кремния $N_c = 6$.

Мы видим, что величина m^* в (IV.3.27) заменяется на $N_c^{2/3} (m_1 m_2 m_3)^{1/3}$.

Для того чтобы (2.22) имело вид, аналогичный (IV.3.27), можно ввести «эффективную массу для плотности состояний»

$$m_{\text{eff}} = N_c^{2/3} (m_1 m_2 m_3)^{1/3}. \qquad (2.22\text{а})$$

Таким образом, при вычислении химического потенциала из (2.12б) надо заменить m^* на m_{eff}.

Для дырок в валентной зоне германия и кремния энергия для каждой из «гофрированных» поверхностей может быть приведена к виду (IV.15.5)

$$\varepsilon = \frac{\hbar^2 k^2}{2m} \Phi(\vartheta, \varphi), \qquad (2.23)$$

где m — масса электрона и $\Phi(\vartheta, \varphi)$ — определенная функция полярных углов ϑ и φ. При определении $g(\varepsilon)$ нас по сути дела будет интересовать переход от переменных k, ϑ, φ к ε, ϑ, φ, т. е. от k к ε при постоянных ϑ и φ. Число состояний в 1 $см^3$ в объеме $k^2\, dk \sin\vartheta\, d\vartheta\, d\varphi = k^2\, dk\, d\Omega$ равно

$$\frac{k^2\, dk\, d\Omega}{8\pi^3} = \frac{\sqrt{2}}{2\pi^2} \frac{m^{3/2}}{\hbar^3} \sqrt{\varepsilon}\, d\varepsilon \frac{1}{4\pi} \Phi^{-3/2}(\vartheta, \varphi)\, d\Omega,$$

где мы использовали связь между k и ε (2.23) при постоянных ϑ и φ. Отсюда

$$g(\varepsilon) = \frac{\sqrt{2}}{2\pi^2} \frac{m^{3/2}}{\hbar^3} \sqrt{\varepsilon} \sum_n \frac{1}{4\pi} \int \Phi_n^{-3/2}(\vartheta, \varphi)\, d\Omega, \qquad (2.24)$$

где интегрирование ведется по полному телесному углу, а суммирование — по разным энергетическим поверхностям (тяжелым и легким дыркам).

Для дырок «эффективная масса для плотности состояний»

$$m_{\text{eff}} = m \left[\sum_n \frac{1}{4\pi} \int \Phi_n^{-3/2}(\vartheta, \varphi)\, d\Omega \right]^{2/3}. \qquad (2.24\text{а})$$

В табл. VI.1 приведены отношения $\left(\frac{m_{\text{eff}}}{m}\right)^{3/2}$ для электронов и дырок в германии и кремнии. Произведение $m^{3/2}$ на число, взятое из таблицы, дает $m_{\text{eff}}^{3/2}$, непосредственно входящее в плотность состояний.

6. Согласно уравнениям (2.16б) или (2.18б), при понижении температуры до абсолютного нуля ($T = 0$) концентрация электронов также стремится к нулю и, следовательно, сопротивление полупроводника неограниченно возрастает. Однако сопротивление большинства полупроводников при $T \to 0$ остается конечным.

Причина этого заключается в том, что при достаточно большой плотности примесных центров волновые функции локализованных электронов перекрываются. При этом энергетические уровни этих электронов несколько расширяются — образуется

Таблица VI.1

Носители тока	Ge (4минимума)	Si (6 минимумов)
Электроны	0,412	1,129
Дырки (1)	0,208	0,390
Дырки (2)	0,0084	0,068
Все дырки	0,216	0,458
Среднее геометрическое значение $\frac{m_{\text{eff}}^{3/2}}{m^{3/2}}$ для дырок и электронов	0,299	0,719

примесная зона и электроны получают возможность передвигаться по всему объему кристалла. Такой механизм электрического тока получил название *примесной проводимости*. Обычно подвижность электронов в примесной зоне намного меньше подвижности электронов в зоне проводимости. Примесная подвижность слабо зависит от температуры и быстро уменьшается по мере уменьшения концентрации примесей. Однако она наблюдается и при сравнительно малых концентрациях примесных центров, например, в германии — до концентраций порядка 10^{15} $см^{-3}$.

При более высоких концентрациях примесей примесная зона сливается с зоной проводимости, при этом электроны остаются в ней при всех температурах и их концентрация не зависит от температуры. Такие вещества называются *полуметаллами*. Перекрытие волновых функций электронов на примесях наступает тем легче, чем больше их боровский радиус, т. е. чем больше диэлектрическая постоянная ε и меньше эффективная масса электрона m^*. Поэтому вещества с большим ε и малой m^* становятся полуметаллами при сравнительно низких концентрациях примеси; так, в n-In Sb уже при $n_D \approx 10^{15}$ $см^{-3}$ все примеси остаются ионизованными при любой температуре, а в n-Ge такой эффект наблюдается только при концентрациях, больших $5 \cdot 10^{18}$ $см^{-3}$. В очень чистых образцах, например в n- и p-германии с концентрацией примесей 10^{13} $см^{-3}$, проводимость по примесной зоне не наблюдается и их сопротивление при понижении температуры возрастает в соответствии с уравнением (2.18б).

§ 3. Теплоемкость свободных электронов в металлах и полупроводниках

1. При изучении теплоемкости кристаллической решетки (гл. III, § 11) мы уже отмечали то обстоятельство, что металлы при высоких температурах подчиняются закону Дюлонга и Пти

и что это находится в кажущемся противоречии с тем, что, согласно классической статистике, свободные электроны должны участвовать в теплоемкости тела. Разрешение этого противоречия заключается в том, что при тех температурах, при которых колебания решетки подчиняются законам классической статистики, свободные электроны металла вырождены и подчиняются статистике Ферми—Дирака.

Рассмотрим теплоемкость свободных электронов металла и их вклад в общую теплоемкость кристалла.

Энергия свободных электронов металла в 1 $см^3$ равна

$$\mathscr{E} = 2\int_0^\infty \varepsilon f_0(\varepsilon)\, g(\varepsilon)\, d\varepsilon. \tag{3.1}$$

Вычислим $\mathscr{E}$ с точностью до $(k_0T/\zeta_0)^2$. Для этого надо воспользоваться формулой (2.7а). В нашем случае $\varphi(\varepsilon) = \frac{1}{5\pi^2}\left[\frac{(2m^*)^{3/2}}{\hbar^3}\right]\varepsilon^{5/2}$. Применяя общий метод, развитый в Приложении 17, получим

$$\mathscr{E} = \frac{(2m^*)^{3/2}}{5\pi^2\hbar^3}\zeta^{5/2}\left[1+\frac{5\pi^2}{8}\left(\frac{k_0T}{\zeta}\right)^2\right]. \tag{3.2}$$

Согласно (2.9), ζ само зависит от температуры, так что с интересующей нас точностью

$$\zeta^{5/2} = \zeta_0^{5/2}\left[1-\frac{\pi^2}{12}\left(\frac{k_0T}{\zeta_0}\right)^2\right]^{5/2} \approx \zeta_0^{5/2}\left[1-\frac{5\pi^2}{24}\left(\frac{k_0T}{\zeta_0}\right)^2\right], \tag{3.3}$$

где $[\ ]^{5/2}$ разложена по биному Ньютона.

Подставляя (3.3) в (3.2) (в квадратной скобке в (3.2) можно положить $\zeta = \zeta_0$), получим, пренебрегая членами порядка $(k_0T/\zeta_0)^4$:

$$\mathscr{E} = \frac{(2m^*)^{3/2}}{5\pi^2\hbar^3}\zeta_0^{5/2}\left[1+\frac{5\pi^2}{12}\left(\frac{k_0T}{\zeta_0}\right)^2\right]. \tag{3.4}$$

Теплоемкость электронов в 1 $см^3$ при постоянном объеме равна

$$c_V = \frac{\partial\mathscr{E}}{\partial T} = \frac{\sqrt{2}}{3}\frac{m^{*3/2}}{\hbar^3}\zeta_0^{1/2}k_0^2T, \tag{3.5}$$

т. е. пропорциональна T.

При абсолютном нуле температуры энергия электронов

$$\mathscr{E}_0 = \frac{(2m^*)^{3/2}}{5\pi^2\hbar^3}\zeta_0^{5/2} = \frac{3^{5/3}\pi^{4/3}\hbar^2}{10m^*}n^{5/3}, \tag{3.4а}$$

что совпадает с (П.4.5). Теплоемкость при абсолютном нуле $c_V^{(0)} = \left(\frac{\partial\mathscr{E}}{\partial T}\right)_0 = 0$, в соответствии с теоремой Нернста.

Выражение (3.5) имеет наглядный смысл: в тепловом возбуждении участвуют только электроны в узком слое энергии k_0T

вблизи энергии ζ_0 и их число порядка $g(\zeta_0) k_0 T$; каждый из этих электронов вносит в теплоемкость вклад порядка k_0, поэтому их полная теплоемкость $\sim g(\zeta_0) k_0^2 T$, что по порядку величины совпадает с (3.5).

Оценим отношение теплоемкости электронов (3.5) к теплоемкости 1 $см^3$ решетки. Так как в металле концентрация электронов n порядка числа атомов решетки в 1 $см^3$, то теплоемкость 1 $см^3$ решетки при высоких температурах порядка nk_0, а интересующее нас отношение

$$\frac{m^{*3/2} \zeta_0^{1/2} k_0 T}{\hbar^3 n} \approx \frac{k_0 T}{\zeta_0}, \tag{3.6}$$

где мы n подставили из (2.8). Так как ζ_0 порядка нескольких электрон-вольт, а $k_0 T$ при комнатных температурах $\approx 0{,}03$ $эв$, то отношение (3.6) порядка долей процента. Этим объясняется практическое неучастие свободных электронов в теплоемкости металла в широкой области температур.

При низких температурах теплоемкость решетки убывает как T^3 (III.11.17), т. е. быстрее (3.5), поэтому ниже некоторой температуры T_0 теплоемкость свободных электронов больше теплоемкости решетки. Температуру T_0 можно определить, приравнивая теплоемкость электронов (3.5) теплоемкости решетки (III.11.17) (если отнести ее к 1 $см^3$ кристалла).

2. Теплоемкость невырожденного электронного газа, концентрация которого n не зависит от температуры, может быть вычислена, если в формулу (3.1) подставить вместо $f_0(\varepsilon)$ функцию распределения Максвелла—Больцмана (2.11а). Простой расчет приводит в этом случае к хорошо известному результату:

$$c_V = {}^3/_2 nk_0. \tag{3.7}$$

Интереснее вычислить теплоемкость свободных электронов (дырок) в случае невырожденного полупроводника, когда их число зависит от температуры. На первый взгляд кажется, что если энергия активации электронов порядка $k_0 T$, то электронная теплоемкость может быть сравнима с теплоемкостью решетки. Рассмотрим для примера собственный полупроводник с шириной запрещенной зоны ε_G. Энергия электронов и дырок в 1 $см^3$ равна

$$\mathscr{E} = 2\int_0^\infty \varepsilon f_0(\varepsilon) g(\varepsilon)\, d\varepsilon + 2\int_0^\infty (\varepsilon' + \varepsilon_G) f_0'(\varepsilon') g(\varepsilon')\, d\varepsilon', \tag{3.8}$$

где первый интеграл относится к электронам, второй—к дыркам. Заменяя $f_0(\varepsilon)$ и $f_0'(\varepsilon')$ классическими функциями распределения,

а $g(\varepsilon)$ и $g(\varepsilon')$ выражением (2.4), получим

$$\mathscr{E} = \frac{\sqrt{2}}{\pi^2}\frac{m_n^{3/2}}{\hbar^3} e^{\zeta/k_0T} \int_0^\infty \exp\left(-\frac{\varepsilon}{k_0T}\right)\varepsilon^{3/2}\,d\varepsilon + \\ + \frac{\sqrt{2}}{\pi^2}\frac{m_p^{3/2}}{\hbar^3}\exp\left(-\frac{\zeta}{k_0T}\right)\exp\left(-\frac{\varepsilon_G}{k_0T}\right)\int_0^\infty \exp\left(-\frac{\varepsilon'}{k_0T}\right)(\varepsilon'+\varepsilon_G)^{3/2}\,d\varepsilon'. \tag{3.8а}$$

Вводя в интегралах новую переменную $x=\sqrt{\varepsilon}=\sqrt{\varepsilon'}$ и применяя формулу (П.7.3), получим, используя значение химического потенциала (2.16),

$$\mathscr{E} = \frac{(2\pi\sqrt{m_n m_p}k_0T)^{3/2}}{4\pi^3\hbar^3} e^{-\frac{\varepsilon_G}{2k_0T}}(3k_0T+\varepsilon_G) = n(3k_0T+\varepsilon_G), \tag{3.8б}$$

где $n=p$—концентрация электронов или дырок.

Теплоемкость электронов и дырок оказывается равной

$$c_V = \frac{\partial\mathscr{E}}{\partial T} = k_0 n\left[\frac{15}{2}+3\left(\frac{\varepsilon_G}{k_0T}\right)+\frac{1}{2}\left(\frac{\varepsilon_G}{k_0T}\right)^2\right], \tag{3.9}$$

где $n = n_T^0 \exp\left(-\frac{\varepsilon_G}{2k_0T}\right)$ (ср. (2.16б)).

Формула (3.9) справедлива при $\varepsilon_G \geqslant k_0T$ (иначе надо учитывать вырождение свободных электронов и дырок). Легко показать, что наибольшее значение c_V получается при $\varepsilon_G \approx k_0T$; тогда по порядку величины $c_V \approx k_0 n_T^0$.

Теплоемкость 1 *см*³ решетки при температурах выше дебаевской по порядку величины равна $c_V^{(L)} \approx k_0 n_0$, где n_0—число атомов в 1 *см*³. Таким образом, $\frac{c_V}{c_V^{(L)}} \approx \frac{n_T^0}{n_0} \ll 1$.

Аналогичные соотношения имеют место и в случае примесных полупроводников. Таким образом, электронно-дырочная теплоемкость в полупроводниках всегда очень мала по сравнению с теплоемкостью кристаллической решетки.

§ 4. Магнитные свойства вещества. Парамагнетизм газов и электронов проводимости в металлах и полупроводниках

1. Учение о магнитных свойствах веществ (магнетиков) весьма обширно, и мы остановимся только на некоторых основных понятиях [1]).

[1]) Для более подробного ознакомления см. книгу Вонсовского С. В. Магнетизм. Магнитные свойства диа-, пара-, ферро-, антиферро- и ферримагнетиков.— М.: Наука, 1971.

Если поместить тело во внешнее магнитное поле, то каждый элемент его объема dV приобретает магнитный момент

$$d\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{H})\, dV. \tag{4.1}$$

Намагниченность или магнитный момент единицы объема $\boldsymbol{M}$ зависит от величины и направления напряженности магнитного поля $\boldsymbol{H}$. В изотропных телах $\boldsymbol{M}$ параллельно $\boldsymbol{H}$ и не зависит от направления $\boldsymbol{H}$. Для изотропных тел, рассмотрением которых мы в дальнейшем ограничимся, величина

$$\left(\frac{dM}{dH}\right)_{H=0} = \chi \tag{4.1а}$$

называется *магнитной восприимчивостью* вещества. Она зависит от вещества, температуры, давления, но не зависит от магнитного поля $\boldsymbol{H}$.

Если при разложении $\boldsymbol{M}(\boldsymbol{H})$ в ряд по $\boldsymbol{H}$ ограничиться линейным членом, то

$$\boldsymbol{M} = \chi \boldsymbol{H}. \tag{4.1б}$$

Легко увидеть, что χ безразмерна.

Мы различаем:

1) *Парамагнитные вещества*, для которых $\chi > 0$, т. е. $\boldsymbol{M}$ направлено параллельно $\boldsymbol{H}$. Газ кислород парамагнитен; для него $\chi \sim 1/T$ (закон П. Кюри); при комнатной температуре и атмосферном давлении $\chi_{O_2} = 0{,}14 \cdot 10^{-6}$.

Щелочные металлы тоже парамагнитны, но для них χ от температуры практически не зависит; для натрия $\chi_{Na} = 0{,}58 \cdot 10^{-6}$.

2) *Диамагнитные вещества*, для которых $\chi < 0$, т. е. $\boldsymbol{M}$ направлено противоположно полю $\boldsymbol{H}$. Для диамагнетиков χ от температуры зависит слабо (через плотность вещества). Инертные газы диамагнитны; для гелия при нормальных условиях $\chi_{He} = -0{,}02 \cdot 10^{-6}$.

3) *Ферромагнитные вещества* (железо, кобальт, никель и ряд сплавов). Свойства ферромагнетиков очень сложны и в этой книге рассматриваться не будут. Выше некоторой температуры, называемой *температурой Кюри* и равной для Fe, Co, Ni соответственно 1040, 1404 и 631°K, ферромагнитные вещества ведут себя как парамагнетики. Ниже температуры Кюри ферромагнетики обладают рядом интересных особенностей: в чистых монокристаллах уже слабое магнитное поле вызывает в определенных кристаллографических направлениях очень большую намагниченность $\boldsymbol{M}$, величина которой зависит не только от напряженности поля $\boldsymbol{H}$, но и от предыстории образца (*гистерезис*); при уменьшении внешнего магнитного поля до нуля ферромагнетики сохраняют некоторый магнитный момент.

Выведем для магнетиков некоторые простые термодинамические соотношения. Рассмотрим два тела, размеры которых

малы по сравнению с расстоянием между ними R. Если первое тело обладает магнитным моментом $\boldsymbol{M}'$, направленным параллельно R, то оно создает в точке нахождения второго тела магнитное поле [1])

$$H = 2M'/R^3. \tag{4.2}$$

Это поле индуцирует во втором теле магнитный момент M, тоже параллельный R. Сила взаимного притяжения между телами

$$\mathcal{K} = 6MM'/R^4. \tag{4.2а}$$

При увеличении R на dR над системой совершается работа

$$dA = \mathcal{K}\, dR = -M\, dH, \tag{4.2б}$$

где мы воспользовались (4.2а) и (4.2). На самом деле это выражение для dA имеет более общий смысл и не связано с рассматриваемой моделью [2]).

Рассмотрим магнетик фиксированного объема в 1 $см^3$. Из первого начала термодинамики следует, что изменение внутренней энергии системы

$$d\mathcal{E} = dQ + dA, \tag{4.3}$$

где dQ — тепло, обратимо переданное системе, а dA — совершенная над ней работа. Полагая $dQ = T\, dS$ (где dS — изменение энтропии системы) и рассматривая работу dA (4.2б) при изменении магнитного поля, получим

$$d\mathcal{E} = T\, dS - M\, dH. \tag{4.3а}$$

Свободная энергия системы

$$\mathcal{F} = \mathcal{E} - TS. \tag{4.4}$$

Дифференцируя это равенство и используя (4.3а), получим

$$d\mathcal{F} = -S\, dT - M\, dH, \tag{4.4а}$$

откуда

$$\left(\frac{\partial \mathcal{F}}{\partial T}\right)_H = -S, \quad \left(\frac{\partial \mathcal{F}}{\partial H}\right)_T = -M. \tag{4.4б}$$

Отсюда и из (4.1а) следует

$$\chi = -\left(\frac{\partial^2 \mathcal{F}}{\partial H^2}\right)_{T;\ H=0}. \tag{4.5}$$

Таким образом, для определения магнитной восприимчивости χ надо найти выражения либо для магнитного момента $\boldsymbol{M}$ (4.1а), либо для свободной энергии $\mathcal{F}$ (4.5).

---

[1]) Тамм И. Е., гл. II (формулы для напряженности поля в случае электрического и магнитного диполя одинаковы).

[2]) Более общий подход см. Ансельм А. И., с. 92.

2. Рассмотрим теорию намагничения вещества, состоящего из свободно вращающихся частиц (молекул) с постоянными магнитными моментами $\boldsymbol{\mu}_A$ (теория парамагнитного газа Ланжевена). Магнитное поле $\boldsymbol{H}$ ориентирует магнитные моменты $\boldsymbol{\mu}_A$ вдоль поля, тепловое движение дезориентирует их. С классической точки зрения момент $\boldsymbol{\mu}_A$ может быть произвольно ориентирован относительно поля $\boldsymbol{H}$. Потенциальная энергия $\boldsymbol{\mu}_A$ в поле $\boldsymbol{H}$ равна $-(\boldsymbol{\mu}_A\boldsymbol{H}) = -\mu_A H\cos\vartheta$, где ϑ — угол между $\boldsymbol{\mu}_A$ и $\boldsymbol{H}$. Вероятность $\boldsymbol{\mu}_A$ быть направленным к полю $\boldsymbol{H}$ под углом, лежащим между ϑ и $\vartheta+d\vartheta$, равна

$$dw = A\exp\frac{\mu_A H\cos\vartheta}{k_0T}\sin\vartheta\,d\vartheta, \tag{4.6}$$

где $2\pi\sin\vartheta\,d\vartheta = d\Omega$ — телесный угол, соответствующий интервалу угла $d\vartheta$, а A — константа, определяемая из условия нормировки вероятности:

$$\int dw = A\int\limits_0^{\pi}\exp\frac{\mu_A H\cos\vartheta}{k_0T}\sin\vartheta\,d\vartheta = A\int\limits_{-1}^{+1}e^{\alpha x}\,dx = 1, \tag{4.6а}$$

где $\alpha = \mu_A H/k_0T$ и $x = \cos\vartheta$.

Среднее значение проекции момента $\boldsymbol{\mu}_A$ на направление магнитного поля

$$\langle\mu_A\rangle = \int\mu_A\cos\vartheta\,dw = \mu_A\frac{\int\limits_{-1}^{+1}xe^{\alpha x}\,dx}{\int\limits_{-1}^{+1}e^{\alpha x}\,dx}. \tag{4.7}$$

Интеграл в знаменателе берется элементарно, а интеграл числителя равен производной по α от интеграла знаменателя. Таким образом,

$$\langle\mu_A\rangle = \mu_A L\left(\frac{\mu_A H}{k_0T}\right), \tag{4.7а}$$

где функция Ланжевена

$$L(\alpha) = \frac{e^{\alpha}+e^{-\alpha}}{e^{\alpha}-e^{-\alpha}} - \frac{1}{\alpha} = \operatorname{cth}\alpha - \frac{1}{\alpha}. \tag{4.7б}$$

Для слабых полей, когда $\mu_A H \ll k_0T$, т. е. $\alpha \ll 1$, получим, разлагая правую часть (4.7б) в ряд:

$$L(\alpha) \approx \alpha/3. \tag{4.7в}$$

В этом случае намагниченность

$$M = n\langle\mu_A\rangle = \frac{\mu_A^2 n}{3k_0T}H, \tag{4.8}$$

откуда магнитная восприимчивость

$$\chi = M/H = \mu_A^2 n/3k_0 T. \qquad (4.8a)$$

Мы видим, что в согласии с законом П. Кюри $\chi \sim \frac{1}{T}$.

В квантовой теории парамагнетизма атомов и ионов необходимо учесть два важных обстоятельства: дискретность пространственного квантования момента количества движения электрона и наличие у него спи́на.

В квантовой механике показывается [1]), что если электрон в атоме находится в стационарном состоянии с определенной проекцией момента количества движения (вращательного момента) $\mathscr{L}_z = \hbar m$ (m — магнитное квантовое число), то этому состоянию соответствует магнитный момент $M_z = \mu_B m$, где

$$\mu_B = \frac{e\hbar}{2mc} = 0{,}927 \cdot 10^{-20} \text{ эрг/э} \qquad (4.9)$$

— *магнетон Бора* — элементарный (наименьший) магнитный момент, фигурирующий в квантовой теории.

Таким образом, для орбитального движения электрона

$$\frac{M_z}{L_z} = \frac{e}{2mc}. \qquad (4.9a)$$

Такое же значение имеет отношение этих величин в классической теории (см. в следующем параграфе формулу (5.2)).

Теория и опыт показывают, что свободный электрон обладает магнитным моментом, равным магнетону Бора μ_B, и вращательным моментом, проекции которого на заданное направление равны $s_z = \pm \hbar/2$. Эти свойства электрона получили название *спи́на*.

Для спина электрона отношение $\mu_B/|s_z|$ имеет «аномальное», вдвое большее значение, чем в (4.9a).

Складывая векторно [2]) магнитные и вращательные моменты орбитального движения и спина, легко понять, что в силу «аномальности» отношения $\mu_B/|s_z|$ для спина, направление результирующего магнитного момента не будет совпадать с направлением результирующего вращательного момента.

Это обстоятельство служит причиной аномального эффекта Зеемана. Можно показать [3]), что в результате прецессии результирующего магнитного момента вокруг результирующего вращательного момента в направлении последнего возникает эффективный магнитный момент

$$\mu_j = \mu_B j g, \qquad (4.10)$$

---

[1]) Блохинцев Д. И., § 53.

[2]) Так называемая векторная модель атома обосновывается в квантовой механике посредством теории групп.

[3]) Блохинцев Д. И., § 74.

где j—квантовое число полного вращательного момента, равного $\hbar\sqrt{j(j+1)}$, а

$$g=1+\frac{j(j+1)+s(s+1)-l(l+1)}{2j(j+1)} \tag{4.10a}$$

—множитель Ландэ. Здесь l—квантовое число орбитального вращательного момента, равного $\hbar\sqrt{l(l+1)}$, а s—спиновое квантовое число принимающее два значения: $+1/2$ и $-1/2$.

Во внешнем магнитном поле $\boldsymbol{H}$ результирующий вращательный момент электрона в атоме принимает $2j+1$ дискретных ориентаций, образуя с магнитным полем углы: $\cos(\widehat{\boldsymbol{H}, j})=m/j$, где m—магнитное квантовое число, принимающее значения $j, j-1, \ldots, 0, \ldots, -j+1, -j$. Энергия магнитного момента μ_j в магнитном поле равна

$$u_\mu=-\mu_j H\cos(\widehat{\boldsymbol{H}, j})=-\mu_B H g m. \tag{4.11}$$

Среднее значение магнитного момента в направлении поля равно

$$\langle\mu\rangle=\frac{\sum_{m=-j}^{+j}\mu_j\cos(\widehat{\boldsymbol{H}, j})\,e^{-\frac{u_\mu}{k_0T}}}{\sum_{m=-j}^{+j}e^{-\frac{u_\mu}{k_0T}}}=\mu_B g\frac{\sum_{m=-j}^{+j}me^{\alpha m}}{\sum_{m=-j}^{+j}e^{\alpha m}}, \tag{4.12}$$

где $\alpha=\mu_B g H/k_0T$.

Рассмотрим для простоты случай слабых магнитных полей, когда $\alpha\ll 1$, так что можно положить $\exp(\alpha m)\approx 1+\alpha m$. В этом случае из (4.12) легко следует

$$\langle\mu\rangle=\frac{\mu_B^2 g^2 j(j+1)}{3k_0T}H, \tag{4.13}$$

так что магнитная восприимчивость аналогично (4.8а) равна

$$\chi=n\langle\mu\rangle/H=\mu_{\text{eff}}^2 n/3k_0T, \tag{4.13a}$$

где эффективный магнитный момент $\mu_{\text{eff}}=\mu_B g\sqrt{j(j+1)}$. Мы видим, что и в квантовом случае χ обратно пропорционально T.

3. Рассмотрим парамагнетизм свободных электронов (электронного газа), связанный с существованием магнитного момента у электрона (спина). В этом случае орбитальный вращательный момент равен нулю ($l=0$) и, следовательно, $j=s=1/2$.

Таким образом, множитель Ландэ для свободных электронов

$$g=2, \tag{4.14}$$

как это следует из (4.10а).

Парамагнитная восприимчивость должна была бы, согласно (4.13а), равняться

$$\chi=\mu_B^2 n/k_0T. \tag{4.14a}$$

что должно было бы приводить к значительному и сильно зависящему от температуры парамагнетизму для свободных электронов металла. Однако опыт, как упоминалось выше, показывает, что парамагнитная восприимчивость металлов мала и от температуры почти не зависит. Объяснение этому было дано Паули (1927) на основе представления об электронах проводимости, как о сильно вырожденном ферми-газе. Работа эта положила начало квантовой теории металлов. Из квантовой механики известно, что при наличии внешнего магнитного поля H магнитный момент электрона может быть направлен либо по полю, и тогда его энергия равна $\varepsilon - \mu_B H$, либо против поля, и тогда его энергия равна $\varepsilon + \mu_B H$, где ε — энергия электрона без поля.

Очевидно, что суммарный магнитный момент 1 $см^3$ вещества, обусловленный электронами проводимости с моментами, направленными по полю, равен

$$M_+ = \mu_B \int f_0(\varepsilon - \mu_B H)\, g(\varepsilon)\, d\varepsilon,$$

так как $g(\varepsilon)\, d\varepsilon$ — число квантовых состояний в интервале $d\varepsilon$ со спином одного направления.

Аналогично суммарный магнитный момент, обусловленный электронами с моментами, направленными против поля

$$M_- = \mu_B \int f_0(\varepsilon + \mu_B H)\, g(\varepsilon)\, d\varepsilon.$$

Таким образом, результирующий момент

$$M = M_+ - M_- = \mu_B \int \{f_0(\varepsilon - \mu_B H) - f_0(\varepsilon + \mu_B H)\}\, g(\varepsilon)\, d\varepsilon. \quad (4.15)$$

Для малых магнитных полей можно разложить $f_0(\varepsilon \mp \mu_B H)$ в ряд по $\mu_B H$ и ограничиться членом первой степени; тогда

$$M = H\mu_B^2 2 \int \left(-\frac{\partial f_0(\varepsilon)}{\partial \varepsilon}\right) g(\varepsilon)\, d\varepsilon. \quad (4.15а)$$

Легко видеть, что $\frac{\partial f_0}{\partial \varepsilon} = -\frac{\partial f_0}{\partial \zeta}$; поэтому

$$\chi = \frac{M}{H} = \mu_B^2 \frac{\partial}{\partial \zeta}\left(2 \int f_0 g\, d\varepsilon\right) = \mu_B^2 \frac{\partial n}{\partial \zeta}, \quad (4.15б)$$

где n — концентрация свободных зарядов.

В случае сильно вырожденных электронов проводимости металла интеграл (4.15а) может быть вычислен по формуле (П.17.5), причем в нашем случае

$$\psi(\eta) = \frac{\sqrt{2}}{\pi^2} \frac{m^{*3/2}}{\hbar^3} (\zeta + k_0 T\eta)^{1/2}.$$

Таким образом,

$$M = H\mu_B^2 \left[\psi(0) + \frac{\pi^2}{6}\psi''(0)\right] = H\mu_B^2 \frac{\sqrt{2}}{\pi^2} \frac{m^{*3/2}}{\hbar^3} \zeta^{1/2} \left[1 - \frac{\pi^2}{24}\left(\frac{k_0 T}{\zeta}\right)^2\right].$$

Заменяя ζ по формуле (2.9), получим с точностью до $(k_0T/\zeta_0)^2$

$$M = H\mu_B^2 g(\zeta_0)\left[1-\frac{\pi^2}{12}\left(\frac{k_0T}{\zeta_0}\right)^2\right]. \qquad (4.16)$$

Используя для ζ_0 выражение (2.8), получим

$$\chi = \frac{3^{1/3}}{\pi^{4/3}}\frac{\mu_B^2 m^* n^{1/3}}{\hbar^2}\left[1-\frac{\pi^2}{12}\left(\frac{k_0T}{\zeta_0}\right)^2\right]. \qquad (4.16а)$$

Мы видим, что главная часть парамагнитной восприимчивости электронов проводимости металла, в соответствии с опытом, от температуры не зависит; малая температурная поправка к χ при комнатных температурах порядка $(k_0T/\zeta_0)^2 \approx 10^{-4}$. Сравнивая (4.16) с (4.8а) или (4.14а), видим, что тепловая энергия молекул k_0T в формуле Ланжевена заменяется в случае сильно вырожденного ферми-газа квантованной энергией $\hbar^2/m^*d^2$, где $d = n^{-1/3}$ — среднее расстояние между частицами.

Для невырожденных электронов проводимости полупроводника

$$\frac{\partial n}{\partial \zeta} = \frac{\partial}{\partial \zeta} e^{\frac{(\zeta-\varepsilon)}{k_0T}} = \frac{n}{k_0T}, \qquad (4.17)$$

поэтому из (4.15б) следует (4.14а).

4. За последние годы значительный интерес для изучения свойств твердого тела приобрело исследование парамагнитного резонанса электронов проводимости и электронов примесных центров (Е. К. Завойский, 1946).

Согласно (4.11) уровень энергии атомного электрона в магнитном поле расщепляется на эквидистантные подуровни, отстоящие друг от друга на величину $g\mu_B H$, где g — множитель Ландэ. Для электрона примесного центра или электрона проводимости в кристалле можно также положить расщепление равным $g\mu_B H$, где множитель g, называемый теперь *спектроскопическим множителем расщепления*, учитывает в общем случае как наличие орбитального вращательного момента у электрона, так и его взаимодействие с решеткой. Поэтому для электронов проводимости в кристалле g отлично от 2 и может быть анизотропно, т. е. иметь различные значения для разных ориентаций магнитного поля.

Энергетический уровень электрона проводимости из-за наличия у него спина ($s = \pm 1/2$) расщепляется на два подуровня. Однако и при мультиплетном расщеплении энергетического уровня электрона в магнитном поле правила отбора для дипольного магнитного перехода разрешают переходы только между соседними подуровнями ($\Delta m = \pm 1$). Поэтому резонансное поглощение высокочастотных радиоволн происходит при частоте ω, удовлетворяющей условию

$$\hbar\omega = g\mu_B H. \qquad (4.18)$$

Таким образом, измерение положения резонансных пиков позволяет определить множитель g, знание которого дает возможность сделать ряд заключений о состоянии электрона в кристалле. Измерение ширины и формы резонансного пика позволяет исследовать взаимодействие примесного электрона с магнитными моментами атомов окружающих узлов, взаимодействие электрона с колебаниями решетки и т. д.

Аналогичная картина парамагнитного резонанса наблюдается для атомных ядер, в частности, когда они входят в состав кристаллической решетки.

Рассмотрим полуфеноменологическую теорию парамагнитного резонанса (Блох, 1946).

Если $\boldsymbol{M}$ — вектор намагничения, т. е. магнитный момент единицы объема вещества, а $\boldsymbol{L}$ — результирующий момент количества движения электронов единицы объема, то

$$\frac{d\boldsymbol{L}}{dt} = [\boldsymbol{M}, \boldsymbol{H}], \tag{4.19}$$

что может быть получено суммированием уравнений движения для отдельных частиц в единице объема образца.

Так как на основании наиболее общих соображений, следующих из квантовой механики,

$$\boldsymbol{M} = \gamma \boldsymbol{L}, \tag{4.20}$$

где

$$\gamma = eg/2mc, \tag{4.20a}$$

то

$$\frac{d\boldsymbol{M}}{dt} = [\gamma \boldsymbol{M}, \boldsymbol{H}]. \tag{4.21}$$

Пусть большое постоянное магнитное поле направлено вдоль оси z (так, что $H_z \equiv H_0$) и малое переменное высокочастотное магнитное поле — вдоль оси x: $H_x = H_1 \exp(i\omega t)$ $(H_1 \ll H_0)$. Ищем решение (4.21) в виде

$$M_x = M_{1x} e^{i\omega t}, \quad M_y = M_{1y} e^{i\omega t}, \quad M_z = M_{0z} + M_{1z} e^{i\omega t}. \tag{4.22}$$

При этом M_{1x}, M_{1y}, M_{1z} порядка H_1 и M_{0z} порядка H_0.

Подставляя (4.22) в (4.21), получим

$$\begin{aligned} &i\omega M_x = \gamma M_y H_z, \\ &i\omega M_y = \gamma (M_z H_x - M_x H_z) \approx \gamma (M_{0z} H_x - M_x H_z), \\ &\frac{d}{dt} M_z = i\omega M_{1z} e^{i\omega t} = -\gamma M_y H_x \approx 0, \end{aligned} \tag{4.23}$$

если пренебречь величинами второго порядка малости по H_1. Таким образом, последнее уравнение можно не учитывать.

Исключая M_y из первых двух уравнений (4.23), получим

$$\chi_x = \frac{M_x}{H_x} = \frac{\chi_0}{1 - (\omega/\omega_0)^2}, \tag{4.24}$$

где

$$\chi_0 = \frac{M_{0z}}{H_z} \text{ и } \omega_0 = \gamma H_0. \tag{4.24а}$$

Так как мы не учли «сил трения», т. е. процессов релаксации прецессионного движения $\boldsymbol{M}$, то амплитуда «переменной восприимчивости» χ_x при резонансной частоте $\omega \approx \omega_0$ стремится к бесконечности.

§ 5. Диамагнетизм атомов и электронов проводимости. Магнитные свойства полупроводников

1. В предыдущем параграфе мы отметили, что существуют вещества (диамагнитные), для которых магнитная восприимчивость $\chi < 0$, т. е. индуцированный магнитный момент направлен против поля $\boldsymbol{H}$. В этом параграфе мы покажем, что квантовые системы движущихся зарядов (электроны атомов и молекул, электроны проводимости) всегда обладают диамагнитными свойствами, которые в некоторых случаях маскируются более сильным парамагнетизмом.

Рассмотрим вначале отдельный атом с Z электронами, помещенный во внешнее магнитное поле $\boldsymbol{H}$. Начало прямоугольной координатной системы поместим в ядро атома, а ось z направим по магнитному полю $\boldsymbol{H}$. Согласно известной теореме Лармора [1]) действие магнитного поля на электроны сводится в первом приближении к равномерному вращению всей системы электронов вокруг оси z (поля $\boldsymbol{H}$) с постоянной угловой скоростью

$$\omega_L = eH/2mc, \tag{5.1}$$

называемой *ларморовой частотой*. Здесь e и m — заряд и масса электрона, c — скорость света. Если смотреть в направлении оси $+z$ (т. е. поля $\boldsymbol{H}$), то вращение электронов происходит по часовой стрелке и, следовательно, магнитный момент μ_0, связанный с соответствующим током, направлен против магнитного поля $\boldsymbol{H}$. Между μ_0 и механическим моментом количества движения электронов l существует универсальное соотношение [2])

$$\mu_0 = \frac{e}{2mc} l. \tag{5.2}$$

Так как

$$l = m\omega_L \sum_{i=1}^{Z} \overline{(x_i^2 + y_i^2)}, \tag{5.3}$$

[1]) Калашников С. Г. Электричество. — 4 изд. — М.: Наука, 1977, § 115.

[2]) См. там же.

где $\overline{x_i^2+y_i^2}$ — сумма квадратов координат i-го электрона, усредненная по объему атома, то

$$\mu_0 = \frac{e^2H}{4mc^2}\sum_{i=1}^{Z}\overline{(x_i^2+y_i^2)}. \tag{5.4}$$

Диамагнитная восприимчивость газа таких атомов

$$\chi_{\text{diam}} = \frac{n\mu_0}{H} = \frac{e^2n}{4mc^2}\sum_{i=1}^{Z}\overline{(x_i^2+y_i^2)}, \tag{5.5}$$

где n — число атомов в 1 $см^3$.

Величина $\sum_{i=1}^{Z}\overline{x_i^2+y_i^2}$ может быть вычислена квантовомеханически; по порядку величины она равна Za^2, где a — линейные размеры атома.

Если атом обладает постоянным магнитным моментом, равным магнетону Бора (4.9), то парамагнитная восприимчивость $\chi_{\text{par}} = \mu_B^2 n/3k_0T$. Отношение диамагнитной восприимчивости (5.5) к парамагнитной по порядку величины равно

$$\frac{\chi_{\text{diam}}}{\chi_{\text{par}}} \approx \frac{e^2nZa^2}{mc^2} : \frac{\mu_B^2 n}{k_0T} = Zk_0T : \frac{e^2}{a}, \tag{5.6}$$

где один множитель a заменен боровским радиусом $\hbar^2/me^2$.

Так как e^2/a — энергия порядка атомной, то при не слишком большом атомном номере Z отношение (5.6) при комнатной температуре порядка 10^{-2}. Вот почему диамагнитные свойства проявляются только в тех веществах, атомы которых не обладают постоянным магнитным моментом.

2. На первый взгляд аналогичный расчет диамагнетизма возможен и для свободных электронов (электронов проводимости). Под действием магнитного поля свободный электрон, обладающий скоростью v, будет описывать в плоскости, перпендикулярной к магнитному полю, круговую орбиту с радиусом

$$r = mcv_\perp/eH = v_\perp/\omega_c, \tag{5.7}$$

где $v_\perp$ — составляющая скорости электрона в плоскости, перпендикулярной к магнитному полю, а

$$\omega_c = eH/mc \tag{5.7a}$$

— так называемая *циклотронная частота*, равная $2\omega_L$, где ω_L — ларморова частота (5.1). Этому круговому движению электрона, согласно (5.2), соответствует магнитный момент

$$\mu = \frac{erv_\perp}{2c} = \frac{\frac{mv_\perp^2}{2}}{H}, \tag{5.8}$$

где мы исключили r посредством равенства (5.7)[1]). Для электронного газа, подчиняющегося классической статистике $\overline{mv_\perp^2}/2 = k_0T$; поэтому диамагнитная восприимчивость

$$\chi = n\frac{k_0T}{H^2}. \tag{5.9}$$

Это выражение не может быть верным, так как оно не зависит от заряда электрона и обратно пропорционально H^2. Можно показать (Бор, 1911), что полученный нами неправильный результат связан с тем, что мы не рассматриваем поверхности (границы) тела, внутри которого движутся электроны. Вблизи поверхности электроны не могут описывать полные круговые орбиты и это приводит к появлению поверхностного тока в образце, который полностью компенсирует магнитные моменты круговых орбит электронов внутри тела.

Покажем в общем виде, что диамагнитная восприимчивость идеального электронного газа, подчиняющегося классической статистике, равна нулю. Положим магнитное поле[2])

$$\boldsymbol{H} = \operatorname{rot}\boldsymbol{A}, \tag{5.10}$$

т. е.

$$H_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z};\quad H_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x};\quad H_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}, \tag{5.10а}$$

где $\boldsymbol{A}(x, y, z)$—вектор-потенциал, определенный с точностью до градиентного преобразования.

Пользуясь формулами (5.10а), легко показать, что для магнитного поля, направленного вдоль оси z ($H_x = H_y = 0$, $H_z = H$), вектор-потенциал можно выбрать в следующем виде:

$$A_x = 0,\quad A_y = xH,\quad A_z = 0. \tag{5.11}$$

Функция Гамильтона для электрона (с зарядом $-e$) в магнитном поле равна[3])

$$\mathcal{H} = \frac{1}{2m}\left(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}\right)^2 + \mathcal{U}(\boldsymbol{r}) = \frac{1}{2m}\left[\left(p_x + \frac{e}{c}A_x\right)^2 + \right.$$
$$\left. + \left(p_y + \frac{e}{c}A_y\right)^2 + \left(p_z + \frac{e}{c}A_z\right)^2\right] + \mathcal{U}(x, y, z), \tag{5.12}$$

где $\mathcal{U}(\boldsymbol{r})$—потенциальная энергия электрона.

Свободная энергия идеального газа, подчиняющегося классической статистике, равна[4])

$$\mathcal{F} = -k_0T\ln Z. \tag{5.13}$$

---

[1]) Мы исключаем r, а не $v_\perp$, так как считаем скорость v заданной.

[2]) Тамм И. Е., Основы теории электричества, М.: Наука, 1976, § 46.

[3]) Ландау Л. Д., Лифшиц Е. М. Теория поля.—3 изд.—М.: Наука, 1973, § 16.

[4]) Ансельм А. И. Основы статистической физики и термодинамика.—М.: 1973, § 3.

Здесь статистический интеграл

$$Z=\left[\int_{-\infty}^{+\infty}\int dp_x\,dp_y\,dp_z\int_{(V)}\int dx\,dy\,dz\,e^{-\frac{\mathcal{H}}{k_0T}}\right]^N, \quad (5.13a)$$

где $\mathcal{H}$ — функция Гамильтона для одной частицы, N — их число.

Для электронов в магнитном поле функция Гамильтона $\mathcal{H}$ равна (5.12). Статистический интеграл Z равен произведению N одинаковых шестикратных интегралов, так как все N электронов независимы и тождественны.

Заменим переменные интегрирования p_x, p_y и p_z на

$$\pi_x=p_x+\frac{e}{c}A_x;\quad \pi_y=p_y+\frac{e}{c}A_y;\quad \pi_z=p_z+\frac{e}{c}A_z. \quad (5.14)$$

При этом $dp_x\,dp_y\,dp_z=d\pi_x\,d\pi_y\,d\pi_z$; пределы интегрирования остаются прежними, и так как определенный интеграл от переменных интегрирования не зависит, то Z, а следовательно, и свободная энергия $\mathcal{F}$ от магнитного поля H не зависят. В этом случае, согласно (4.5), восприимчивость χ равна нулю.

Совершенно аналогично можно показать, что диамагнитная восприимчивость электронного газа, подчиняющегося статистике Ферми, тоже равна нулю. Таким образом, при квазиклассическом рассмотрении электронный газ не обладает диамагнетизмом.

Может возникнуть вопрос, не находится ли этот результат в противоречии с не равным нулю диамагнетизмом атомов (5.5). На самом деле (5.5) получено нами на основании существенно квантового предположения о стационарном состоянии электронов в атоме, не подчиняющихся законам классической статистики.

3. Л. Д. Ландау (1930) показал, что при квантовомеханическом рассмотрении движения свободных электронов в магнитном поле их диамагнитная восприимчивость равна $^1/_3$ парамагнитной восприимчивости. Этот факт не противоречит предыдущему выводу, так как он связан с квантованием движения свободного электрона в магнитном поле.

Пусть на электрон проводимости, помимо периодического потенциала $V(\boldsymbol{r})$, действует однородное магнитное поле, направленное по оси z. Вектор-потенциал магнитного поля может быть выбран в форме (5.11). Из (5.12) следует, что уравнение Шредингера имеет вид [1])

$$\frac{1}{2m}\left[\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2+\left(\frac{\hbar}{i}\frac{\partial}{\partial y}+\frac{e}{c}Hx\right)^2+\left(\frac{\hbar}{i}\frac{\partial}{\partial z}\right)^2\right]\psi(\boldsymbol{r})+V(\boldsymbol{r})\,\psi(\boldsymbol{r})=\mathcal{E}\psi(\boldsymbol{r}). \quad (5.15)$$

[1]) Уравнение Шредингера получается по известному рецепту: $\hat{\mathcal{H}}\psi=\mathcal{E}\psi$, где $\hat{\mathcal{H}}$ — гамильтониан, получаемый из функции Гамильтона при замене p_x на $\frac{\hbar}{i}\frac{\partial}{\partial x}$ и т. д.

Периодический потенциал кристалла $V(\boldsymbol{r})$ может быть опущен, если заменить массу свободного электрона m на эффективную массу m^*. Тогда вместо (5.15) получим

$$-\frac{\hbar^2}{2m^*}\left[\frac{\partial^2\psi}{\partial x^2}+\left(\frac{\partial}{\partial y}+\frac{ieHx}{\hbar c}\right)^2\psi+\frac{\partial^2\psi}{\partial z^2}\right]=\mathscr{E}\psi. \tag{5.16}$$

Так как это уравнение не содержит в явном виде y и z, то ищем решение в виде

$$\psi(x,\ y,\ z)=\varphi(x)\,e^{i(k_y y+k_z z)}. \tag{5.17}$$

Подставляя (5.17) в (5.16) и сокращая на экспоненциальный множитель, получим после простого преобразования

$$-\frac{\hbar^2}{2m^*}\frac{d^2\varphi}{dx^2}+\frac{1}{2}m^*\omega_0^2(x-x_0)^2\varphi=\mathscr{E}_1\varphi, \tag{5.18}$$

где

$$\omega_0=\frac{eH}{m^*c},\quad x_0=-\frac{\hbar c}{eH}k_y \text{ и } \mathscr{E}_1=\mathscr{E}-\frac{h^2k_z^2}{2m^*}. \tag{5.18а}$$

Мы видим, что движение электрона вдоль координаты x описывается уравнением Шредингера для линейного гармонического осциллятора с массой m^*, собственной частотой $\omega_0=2\omega_L^*$ (ω_L^* — ларморова частота частицы с массой m^*), колеблющегося около положения равновесия x_0. Собственные значения энергии $\mathscr{E}_1$ для осциллятора, аналогично (III.10.4), равны

$$\mathscr{E}_1=\mathscr{E}-\frac{\hbar^2k_z^2}{2m^*}=\left(N+\frac{1}{2}\right)\hbar\omega_0=(2N+1)\,\mu^*H, \tag{5.19}$$

где $\mu^*=\left(\frac{m}{m^*}\right)\mu_B$, а собственные функции уравнения (5.18) имеют вид

$$\varphi(x)=\frac{\exp\left[-\frac{1}{2}\left(\frac{x-x_0}{\lambda}\right)^2\right]}{\sqrt{\lambda}}H_N\left(\frac{x-x_0}{\lambda}\right), \tag{5.19а}$$

где $\lambda\equiv\left(\frac{\hbar}{m^*\omega_0}\right)^{1/2}=\left(\frac{\hbar c}{eH}\right)^{1/2}$ — так называемая «магнитная длина» [1]), а H_N — полином Эрмита N-го порядка.

Выражение (5.17) имеет вид, который на первый взгляд свидетельствует о физической эквивалентности направлений y и z, а не x и y, как это имеет место на самом деле (магнитное поле направлено по оси z). Это обстоятельство связано с тем, что операторы, соответствующие координатам «центра окружности» движущегося электрона (один из этих операторов равен

[1]) Если H выражать в эрстедах, то $\lambda=\frac{10^{-3}}{4\sqrt{H}}$ *см*.

$x_0 = -ep_y/eH$), не коммутируют друг с другом[1]). Более симметричное выражение для волновой функции $\psi(\boldsymbol{r})$, хотя и менее удобное для приложений, можно получить, решая задачу в цилиндрических координатах.

Собственные значения уравнения (5.16)

$$\mathscr{E}(N,\ k_z) = (2N+1)\,\mu^* H + \frac{\hbar^2 k_z^2}{2m^*} \tag{5.19б}$$

вырождены по квантовому числу k_y.

Для того чтобы подсчитать число квантовых состояний электрона $g(\mathscr{E})\,d\mathscr{E}$ на интервале энергии $\mathscr{E}$, $\mathscr{E}+d\mathscr{E}$, наложим на волновую функцию (5.17) условия цикличности по осям y и z, т. е. потребуем, чтобы координатам $y \pm L_2$, $z \pm L_3$ соответствовала та же волновая функция, что и координатам y и z. Легко видеть, что для этого необходимо, чтобы

$$k_y = \frac{2\pi}{L_2} n_2, \quad k_z = \frac{2\pi}{L_3} n_3, \tag{5.20}$$

где n_2 и n_3 — любые целые числа. В самом деле: $\exp[ik_y(y \pm L_2)] = \exp(ik_y y)\exp(\pm i2\pi n_2) = \exp(ik_y y)$.

Пусть толщина тела по оси x равна L_1; при этом мы не будем налагать условий цикличности на волновую функцию вдоль x (это неудобно, так как выражение (5.17) непериодично по x), но будем считать, что решение (5.17) существует только в области

$$0 < |x_0| < L_1, \tag{5.21}$$

так что $|x_0|^{\max} = L_1$.

Число квантовых состояний для заданного значения $\mathscr{E}(N, k_z)$ определяется степенью вырождения по k_y и равно $\left(\frac{L_2}{2\pi}\right) k_y^{\max}$. Из выражений (5.18а) и (5.21) $k_y^{\max} = \left(\frac{eH}{\hbar c}\right)|x_0|^{\max} = \frac{eH}{\hbar c} L_1$, так что интересующая нас степень вырождения равна

$$\frac{eH}{2\pi\hbar c} L_1 L_2. \tag{5.22}$$

Определим теперь число квантовых состояний $Z(\mathscr{E})$, энергия которых меньше $\mathscr{E}$. Число квантовых состояний с энергиями от $\mathscr{E}_1 = (2N+1)\,\mu^* H$ до $\mathscr{E}$ равно

$$\frac{2|k_z|L_3}{2\pi} = \frac{2(2m^*)^{1/2} L_3}{2\pi\hbar} \sqrt{\mathscr{E} - (2N+1)\,\mu^* H}, \tag{5.22а}$$

где множитель 2 учитывает два знака у k_z при определении его из (5.19). Число квантовых состояний с энергиями, меньшими $\mathscr{E}$,

[1]) Подробнее об этом см. Ландау Л. Д., Лифшиц Е. М. Квантовая механика. — М., 1974, с. 522.

при всех возможных N равно

$$Z(\mathscr{E}) = \frac{2(2m^*)^{1/2}L_3}{2\pi\hbar}\sum_N \sqrt{\mathscr{E}-(2N+1)\mu^*H}. \qquad (5.22б)$$

Здесь суммирование по N распространяется на все неотрицательные значения подкоренного выражения. Так как степень вырождения каждого состояния (5.22б) равна (5.22), то полное число квантовых состояний с энергией, меньшей $\mathscr{E}$, равно

$$Z(\mathscr{E}) = \frac{2(2m^*)^{1/2}eHV}{(2\pi\hbar)^2c}\sum_N [\mathscr{E}-(2N+1)\mu^*H]^{1/2}, \qquad (5.23)$$

где $V = L_1L_2L_3$ — объем тела.

Плотность числа состояний на $V = 1$ *см*^{3} без учета спина равна

$$g(\mathscr{E}) = \frac{dZ(\mathscr{E})}{d\mathscr{E}} = \frac{(2m^*)^{1/2}eH}{(2\pi\hbar)^2c}\sum_N [\mathscr{E}-(2N+1)\mu^*H]^{-1/2}; \qquad (5.23а)$$

свободная энергия 1 *см*^{3} равна [1])

$$\mathscr{F} = n\zeta - 2k_0T\int^{\infty} \ln\left(1+e^{\frac{(\zeta-\varepsilon)}{k_0T}}\right) g(\mathscr{E})\,d\mathscr{E} =$$

$$= n\zeta - 2k_0T\int^{\infty} \ln\left(1+e^{\frac{(\zeta-\varepsilon)}{k_0T}}\right) \frac{dZ(\mathscr{E})}{d\mathscr{E}}\,d\mathscr{E}. \qquad (5.24)$$

Здесь n — концентрация свободных электронов, множитель 2 учитывает спин; нижний предел интеграла для каждого из слагаемых (5.23а) равен $(2N+1)\mu^*H$.

Интегрируя по частям, получим

$$\mathscr{F} = n\zeta - 2\int^{\infty} f_0(\mathscr{E})\,Z(\mathscr{E})\,d\mathscr{E}, \qquad (5.24а)$$

где $f_0(\mathscr{E})$ — функция распределения Ферми.

Введем новую переменную интегрирования и обозначения:

$$\varepsilon = \frac{\mathscr{E}}{2\mu^*H}, \quad \varepsilon_0 = \frac{\zeta}{2\mu^*H}, \quad \Theta = \frac{k_0T}{2\mu^*H}. \qquad (5.25)$$

Интегрируя (5.24а) еще раз по частям, получим в новых обозначениях:

$$\mathscr{F} = n\zeta + A\int^{\infty} \Phi(\varepsilon)\,\frac{d}{d\varepsilon}\left(\frac{1}{e^{\frac{\varepsilon-\varepsilon_0}{\Theta}}+1}\right)d\varepsilon, \qquad (5.25а)$$

где

$$A = \frac{16m^{3/2}(\mu^*H)^{5/2}}{3\pi^2\hbar^3}, \quad \Phi(\varepsilon) = \sum_N (\varepsilon - N - {}^1/_2)^{3/2} \qquad (5.25б)$$

[1]) Анселъм А. И., формулы (VIII 1.37) и (IX 3.3).

и нижний предел интеграла для каждого слагаемого в $\Phi(\varepsilon)$ определяется из условия вещественности соответствующего радикала.

Используя формулу Пуассона, получим (Приложение 18)

$$\Phi(\varepsilon)=\frac{2}{5}\varepsilon^{5/2}-\frac{1}{16}\varepsilon^{1/2}=$$

$$=\frac{3}{8\pi^2}\sum_{l=1}^{\infty}\frac{(-1)^l}{l^{5/2}}\left[\sin(2\pi l\varepsilon)\cdot S\left(\sqrt{4l\varepsilon}\right)+\cos(2\pi l\varepsilon)\cdot C\left(\sqrt{4l\varepsilon}\right)\right], \quad (5.26)$$

где $S(u)$ и $C(u)$ — интегралы Френеля, записанные в явном виде в (П.18.8а), П.18.8б).

При сильном вырождении электронного газа производная

$$\frac{\partial}{\partial\varepsilon}\left(\frac{1}{\exp\frac{\varepsilon-\varepsilon_0}{\Theta}+1}\right)=-\delta(\varepsilon-\varepsilon_0) \quad (5.27)$$

ведет себя как дельта-функция с эффективной шириной максимума порядка Θ.

Слагаемые в сумме (5.26) с заданным l осциллируют с периодом $\varepsilon=1/l$, т. е. с наибольшим периодом для $l=1$; поэтому, если $\Theta\gg 1$, т. е.

$$k_0T\gg 2\mu^*H, \quad (5.28)$$

то сумма по l в (5.26) даст под знаком интеграла (5.25а) нуль и мы получим

$$\mathcal{F}=n\zeta_0-A\left(\frac{2}{5}\varepsilon_0^{5/2}-\frac{1}{16}\varepsilon_0^{1/2}\right), \quad (5.29)$$

где ζ_0 — энергия Ферми при абсолютном нуле.

Возникает вопрос, не следует ли учитывать зависимость свободной энергии $\mathcal{F}$ от магнитного поля через влияние последнего на химический потенциал ζ. Если химический потенциал $\zeta=\zeta^0+\Delta\zeta$, где ζ^0 — химический потенциал в отсутствие магнитного поля, то

$$\mathcal{F}(\zeta)=\mathcal{F}(\zeta^0+\Delta\zeta)=\mathcal{F}(\zeta^0)+\left(\frac{\partial\mathcal{F}}{\partial\zeta}\right)_0\Delta\zeta+\frac{1}{2}\left(\frac{\partial^2\mathcal{F}}{\partial\zeta^2}\right)_0(\Delta\zeta)^2. \quad (5.30)$$

В состоянии статистического равновесия $(\partial\mathcal{F}/\partial\zeta)_0=0$, поэтому изменение свободной энергии

$$\mathcal{F}(\zeta)-\mathcal{F}(\zeta^0)=\frac{1}{2}\left(\frac{\partial^2\mathcal{F}}{\partial\zeta^2}\right)_0(\Delta\zeta)^2. \quad (5.30а)$$

Так как для изотропного тела изменение химического потенциала $\Delta\zeta$ не зависит от направления магнитного поля, то $\Delta\zeta\sim H^2$ и, следовательно, в наинизшем приближении

$$\mathcal{F}(\zeta)-\mathcal{F}(\zeta^0)\sim H^4. \quad (5.31)$$

Таким образом, в приближении, учитывающим в свободной энер-

гии $\mathcal{F}$ члены порядка H^2, можно считать ζ не зависящим от магнитного поля.

Учитывая значение A (5.25б) и ε_0 (5.25), видим, что и второе слагаемое в (5.29) от магнитного поля не зависит.

Для диамагнитной восприимчивости χ из (5.29) получим

$$\chi = -\frac{\partial^2 \mathcal{F}}{\partial H^2} = -\frac{1}{3}\frac{\sqrt{2}}{\pi^2}\frac{m^{*3/2}\mu^{*2}\zeta_0^{1/2}}{\hbar^3}. \tag{5.32}$$

Если подставить сюда ζ_0 из (2.8а), то непосредственно видно, что по абсолютной величине диамагнитная восприимчивость (5.32) в $\frac{1}{3}\left(\frac{m}{m^*}\right)^2$ раз больше не зависящей от температуры парамагнитной восприимчивости (4.16а). В случае, противоположном (5.28), когда $\Theta \ll 1$, имеем

$$k_0 T \ll 2\mu^* H. \tag{5.33}$$

В этом случае период осцилляций в (5.26) (во всяком случае для $l = 1$) больше эффективной ширины Θ максимума δ-функции в интеграле (5.25а), поэтому в свободной энергии $\mathcal{F}$ появляются осциллирующие члены вида

$$\sin\left(\frac{\pi\zeta_0}{\mu^* H}\right) S\left(\sqrt{\frac{2\zeta_0}{\mu^* H}}\right) + \cos\left(\frac{\pi\zeta_0}{\mu^* H}\right) C\left(\sqrt{\frac{2\zeta_0}{\mu^* H}}\right). \tag{5.34}$$

Очевидно, в магнитной восприимчивости $\chi = -\partial^2\mathcal{F}/\partial H^2$ тоже появляются осциллирующие члены, аналогичные (5.34). Периоды осцилляций тригонометрических функций и интегралов Френеля в (5.34) определяются условием, что $1/H$ меняется на $2\mu^*/\zeta_0$ или на величину порядка $\mu^*/2\zeta_0$.

Подобные осцилляции магнитной восприимчивости были впервые экспериментально обнаружены у висмута и получили название *эффекта де-Гааза—ван-Альфена*, по имени ученых, открывших это явление в 1930 г.

Рассмотрим теперь диамагнитную проницаемость невырожденных электронов проводимости полупроводника. Выражение (5.25а) для свободной энергии остается, конечно, верным и в этом случае, но может быть упрощено: функция распределения Ферми может быть заменена распределением Больцмана $\exp\left(\frac{(\varepsilon_0 - \varepsilon)}{\Theta}\right)$. Таким образом,

$$\mathcal{F} = n\zeta - \frac{Ae^{\varepsilon_0/\Theta}}{\Theta}\sum_{N=0}^{\infty}\int_{N+1/2}^{\infty}\left(\varepsilon - N - \frac{1}{2}\right)^{3/2} e^{-\varepsilon/\Theta}\,d\varepsilon, \tag{5.35}$$

где нижний предел интеграла определяется из условия положительности подкоренного выражения подынтегральной функции.

Введем новую переменную интегрирования:

$$x=\frac{\varepsilon-N-\frac{1}{2}}{\Theta}; \tag{5.36}$$

тогда

$$\mathcal{F}=n\zeta-A\Theta^{3/2}\exp\frac{\zeta}{k_0T}\exp\left(-\frac{1}{2\Theta}\right)\sum_{N=0}^{\infty}e^{-\frac{N}{\Theta}}\int_0^{\infty}x^{3/2}e^{-x}dx. \tag{5.37}$$

Интеграл

$$\int_0^{\infty}x^{3/2}e^{-x}dx=\frac{3\sqrt{\pi}}{4}, \tag{5.38}$$

как это следует из (П.7.12). Бесконечная геометрическая убывающая прогрессия

$$\sum_{N=0}^{\infty}e^{-\frac{N}{\Theta}}=\left(1-e^{-\frac{1}{\Theta}}\right)^{-1}. \tag{5.39}$$

Таким образом,

$$\mathcal{F}=n\zeta-\frac{3\sqrt{\pi}}{4}A\Theta^{5/2}\exp\frac{\zeta}{k_0T}\left[\frac{2ze^{-z}}{1-e^{-2z}}\right], \tag{5.40}$$

где

$$z=1/2\Theta=\mu^*H/k_0T. \tag{5.40а}$$

Так как множитель $A\Theta^{5/2}$ от магнитного поля H не зависит, то зависимость свободной энергии $\mathcal{F}$ от H, если не учитывать слабой зависимости ζ от магнитного поля, определяется квадратной скобкой в (5.40).

В случае невырожденных полупроводников, предельный случай (5.33), т. е. $z\gg1$, особого интереса не представляет[1]). Мы рассмотрим поэтому другой предельный случай (5.28), когда $z\ll1$.

Разлагая в (5.40) квадратную скобку в ряд по степеням z, до членов порядка z^2, получим

$$\left[\frac{2ze^{-z}}{1-e^{-2z}}\right]=\frac{2z}{e^z-e^{-z}}=\frac{z}{\operatorname{sh}z}\approx\frac{z}{z+\frac{z^3}{3!}+\ldots}\approx1-\frac{1}{6}z^2. \tag{5.41}$$

Подставляя это в результат (5.40), получим с точностью до H^2

$$\mathcal{F}=n\zeta-\frac{3\sqrt{\pi}}{4}A\Theta^{5/2}e^{\frac{\zeta}{k_0T}}\left[1-\frac{1}{6}\left(\frac{\mu^*H}{k_0T}\right)^2\right]. \tag{5.42}$$

Используя выражения (5.25), (5.25б) и (2.12а), получим для

[1]) В этом случае $\chi\sim\exp(-z)$, т. е. экспоненциально мала.

диамагнитной восприимчивости

$$\chi = -\frac{\partial^2 \mathcal{F}}{\partial H^2} = -\frac{1}{3}\frac{\mu^{*2} n}{k_0 T} = -\frac{1}{3}\left(\frac{m}{m^*}\right)^2 \frac{\mu_B^2 n}{k_0 T}, \qquad (5.43)$$

равную $\frac{1}{3}\left(\frac{m}{m^*}\right)^2$ парамагнитной восприимчивости (4.14а). Таким образом, соотношение между диа- и парамагнитными восприимчивостями свободных носителей заряда остается одинаковым как для сильно вырожденных электронов металла (5.30), так и для невырожденного электронного газа в полупроводниках (5.43). Можно показать, что это имеет место для любой степени вырождения электронного газа.

4. Магнитная восприимчивость атомного полупроводника

$$\chi = \chi_A + \chi_L + \chi_S + \chi_T, \qquad (5.44)$$

т. е. аддитивно складывается из диамагнитной восприимчивости основной решетки χ_A, пара- и диамагнитной восприимчивости носителей заряда χ_L, восприимчивости примесных центров χ_S и, как показывают новые исследования, восприимчивости термических дефектов (дислокаций, поверхностных уровней, граней и т. д.) [1]).

Дырки аналогично электронам обладают собственным магнитным моментом, равным μ_B (а следовательно, и парамагнитными свойствами). Это следует хотя бы из того, что при удалении электрона из валентной зоны ее магнитный момент меняется на величину μ_B.

Таким образом, невырожденные носители зарядов в смешанном полупроводнике будут обладать восприимчивостью

$$\chi_L = \frac{\mu_B^2 n}{k_0 T}\left[1 - \frac{1}{3}\left(\frac{m}{m_n}\right)^2\right] + \frac{\mu_B^2 p}{k_0 T}\left[1 - \frac{1}{3}\left(\frac{m}{m_p}\right)^2\right], \qquad (5.45)$$

где n и p — концентрации электронов и дырок, а m_n и m_p — их эффективные массы. Если закон дисперсии для энергии электрона (дырки) характеризуется тензором эффективной массы, то отношение m^2/m_n^2, m^2/m_p^2 в (5.45) заменяется отношением m^2 к некоторой простой комбинации компонент тензора эффективных масс.

О величинах χ_A, χ_S и χ_T можно сказать очень мало. Во всяком случае, рассчитать их теоретически с достаточной степенью достоверности в настоящее время невозможно. Поэтому изучение магнитной восприимчивости полупроводников дает сравнительно мало сведений о свойствах носителей тока в них.

---

[1]) Буш Г., Винклер У. Определение характеристических параметров полупроводников. — М., 1959, с. 64.

§ 6. Циклотронный (диамагнитный) резонанс

1. Наиболее полные и достоверные сведения об энергетическом спектре носителей тока (электронов и дырок) в кристаллах были получены при исследовании циклотронного резонанса. В частности, структура энергетических зон германия и кремния, описанная в гл. IV, § 15, в большой мере обоснована изучением в них циклотронного резонанса. Мы сформулируем физические принципы, лежащие в основе явления циклотронного резонанса и дадим его упрощенную теоретическую трактовку.

Квазисвободный электрон (дырка) со скалярной эффективной массой m^* движется в магнитном поле H подобно свободному электрону, описывая в плоскости, перпендикулярной к магнитному полю, круговую орбиту с радиусом (5.7)

$$r = v_\perp/\omega_c, \tag{6.1}$$

где

$$\omega_c = eH/m^*c \tag{6.1а}$$

— *циклотронная частота*, а $v_\perp$ — составляющая скорости в плоскости, перпендикулярной к магнитному полю.

То обстоятельство, что ω_c не зависит от v (в нерелятивистском приближении) было использовано для конструирования ускорителя заряженных частиц — *циклотрона*. Очевидно, что можно попытаться реализовать принцип циклотронного ускорения по отношению к квазисвободным носителям тока в кристалле. Для этого необходимо приложить к образцу, помимо постоянного магнитного поля $\boldsymbol{H}$, высокочастотное электромагнитное поле, колеблющееся в плоскости, перпендикулярной к $\boldsymbol{H}$. Если частота этого поля ω совпадает с ω_c, то электрон будет ускоряться под действием электрического напряжения высокочастотного поля на обоих полуокружностях своего движения. В результате этого его скорость (энергия) будет возрастать и он будет двигаться по раскручивающейся спирали, как заряженная частица в циклотроне. Этот эффект, получивший название *циклотронного (диамагнитного) резонанса* (Дорфман и Дингль, 1951), может быть обнаружен по максимуму поглощения высокочастотного излучения. Определение $\omega_c \approx \omega$ при таком резонансе позволяет наиболее прямым образом определить из (6.1а) эффективную массу m^*.

Помимо взаимодействия с постоянным магнитным и высокочастотным полями, электрон проводимости взаимодействует еще с колебаниями решетки (фононами) или статическими нарушениями решетки (примесными центрами), в результате чего устанавливается некоторое время свободного пробега электрона τ. Приближенно можно себе представить взаимодействие с колебаниями и примесями как некоторую непрерывно прило-

женную к электрону силу трения $\boldsymbol{F}_R$. Если отождествить время свободного пробега τ со временем релаксации скорости электрона $\boldsymbol{v}$ под действием силы трения $\boldsymbol{F}_R$, т. е. считать, что при выключении внешних сил скорость электрона убывает по закону

$$\boldsymbol{v}=\boldsymbol{v}_0 e^{-t/\tau}, \tag{6.2}$$

где $\boldsymbol{v}_0$ — скорость при $t=0$, то сила трения

$$\boldsymbol{F}_R=m^*\frac{d\boldsymbol{v}}{dt}=-\frac{m^*\boldsymbol{v}}{\tau}, \tag{6.2a}$$

т. е. направлена противоположно скорости $\boldsymbol{v}$ и пропорциональна ее величине.

В следующей главе мы более строго выведем время релаксации из кинетического уравнения. При этом выяснится, что в общем случае τ зависит от v. В этой главе мы для простоты будем предполагать, что τ от v не зависит, что не отразится на качественном характере наших выводов.

Для того чтобы можно было наблюдать циклотронный резонанс, необходимо, чтобы

$$\tau \gg 1/\omega_c, \tag{6.3}$$

так как только в этом случае может сформироваться круговая орбита, на которой может реализоваться синхронное ускорение частицы.

Оценим некоторые величины при циклотронном резонансе. Положим частоту переменного электромагнитного поля равной 24 000 *Мгц*. При резонансе $\omega_c/2\pi=24\,000$ *Мгц*, откуда $\omega_c=1{,}5\cdot10^{11}$ *рад/сек*. Для реального отношения $m^*/m\approx0{,}3$ магнитное поле из (6.1а) $H=2\cdot10^3$ *э*. При температуре $T=4°\mathrm{K}$ средняя тепловая скорость электрона $v_T=\left(\frac{8k_0T}{\pi m^*}\right)^{1/2}\approx2{,}4\cdot10^6$ *см/сек* и радиус орбиты (при $v_\perp\approx v_T$) из (6.1) $r\approx5\cdot10^{-5}$ *см*.

2. Рассмотрим элементарную теорию циклотронного резонанса, считая что заряд e обладает скалярной эффективной массой m^* и движется по законам классической механики. Если на заряд действует постоянное магнитное поле $\boldsymbol{H}$, высокочастотное электрическое поле $\boldsymbol{E}$ (действием магнитного поля высокочастотного излучения можно пренебречь) и сила трения $\boldsymbol{F}_R$ (6.2а), то уравнения движения имеют вид

$$m^*\frac{d\boldsymbol{v}}{dt}=-\frac{m^*\boldsymbol{v}}{\tau}+e\left(\boldsymbol{E}+\frac{1}{c}[\boldsymbol{v}\boldsymbol{H}]\right). \tag{6.3a}$$

Предполагая, что $\boldsymbol{H}$ направлено вдоль оси z, $\boldsymbol{E}=\boldsymbol{E}_0\exp(-i\omega t)$ — вдоль оси x и что $\boldsymbol{v}$ пропорционально $\exp(-i\omega t)$, получим в проекциях на оси x и y

$$m^*\left(-i\omega+\frac{1}{\tau}\right)v_x=eE_x+\frac{e}{c}v_yH, \tag{6.4}$$

$$m^*\left(-i\omega+\frac{1}{\tau}\right)v_y=-\frac{e}{c}v_xH. \tag{6.4a}$$

Согласно дифференциальному закону Ома плотность тока

$$j = \sigma E_x = nev_x, \tag{6.5}$$

где σ — удельная электропроводность, n — концентрация носителей тока. Исключая v_y из (6.4) и (6.4а), получим для комплексной электропроводности

$$\sigma = \frac{nev_x}{E_x} = \sigma_0 \frac{1 - i\omega\tau}{1 + (\omega_c^2 - \omega^2)\tau^2 - 2i\omega\tau}, \tag{6.6}$$

где

$$\sigma_0 = ne^2\tau/m^* = ne\mu \tag{6.6а}$$

— удельная электропроводность при $\omega = \omega_c = 0$, т. е. в постоянном электрическом поле $E_x = E_0$ и в отсутствие магнитного поля ($H = 0$). Величина

$$\mu = e\tau/m^* = v_x/E_0, \tag{6.6б}$$

измеряющая среднюю скорость электрона при действии на него постоянного электрического поля с напряженностью, равной 1, называется *подвижностью*.

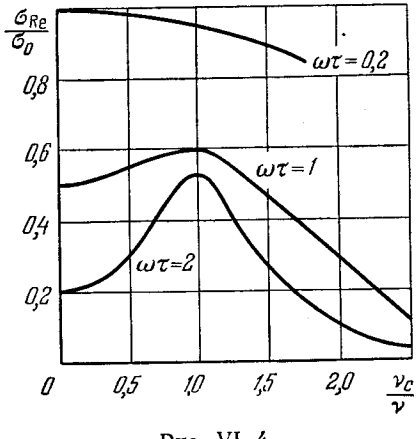


Рис. VI. 4.

Поглощение высокочастотного электромагнитного поля в образце пропорционально вещественной части электропроводности σ (6.6), т. е.

$$\sigma_{\mathrm{Re}} = \sigma_0 \frac{1 + \nu^2 + \nu_c^2}{(1 + \nu_c^2 - \nu^2)^2 + 4\nu^2}, \tag{6.7}$$

где $\nu = \omega\tau$ пропорционально частоте электромагнитного поля и $\nu_c = \omega_c\tau = \mu H/c$ — циклотронной частоте. На рис. VI.4 построены кривые зависимости $\sigma_{\mathrm{Re}}/\sigma_0$ от $\nu_c/\nu = \omega_c/\omega$ при разных значениях $\nu = \omega\tau$. Для $\omega\tau = 2$, когда уже в какой-то мере выполняется условие (6.3), необходимое для наблюдения циклотронного резонанса, кривая обладает ясно выраженным максимумом при $\omega = \omega_c$.

Рассмотрим ряд предельных случаев.

1. $\nu_c \gg \nu$, $\nu_c \gg 1$. Этот случай реализуется в низкочастотном переменном электрическом поле и сильных магнитных полях. Из (6.7) видно

$$\sigma_{\mathrm{Re}} \approx \sigma_0/\nu_c^2 \sim 1/H^2, \tag{6.8}$$

т. е. поглощение убывает обратно пропорционально H^2.

2. $\nu_c \gg \nu$, $\nu_c \ll 1$. Низкочастотное переменное электрическое поле и слабое магнитное поле (либо малое τ — высокие температуры). В этом случае

$$\sigma_{\mathrm{Re}} \approx \sigma_0 (1 - \nu_c^2). \tag{6.9}$$

Если ввести удельное сопротивление $\rho = 1/\sigma$ так, что $\Delta\rho/\rho = -\Delta\sigma/\sigma$, то из (6.9) следует

$$\Delta\rho/\rho_0 = \nu_c^2 = (\mu H/c)^2. \tag{6.9a}$$

Этот результат, как мы увидим в гл. VIII, с точностью до численного множителя совпадает с тем, что получается из кинетического уравнения.

3. $\nu \gg 1$, $\nu_c = 0$. Имеет место облучение высокочастотным полем (например, инфракрасным) в отсутствие магнитного поля. В этом случае коэффициент поглощения пропорционален

$$\sigma_{\mathrm{Re}} \approx \sigma_0/\nu^2. \tag{6.10}$$

4. $\nu = \nu_c \gg 1$. Это условие циклотронного резонанса. Из (6.7) имеем

$$\sigma_{\mathrm{Re}} \approx \frac{1}{2}\sigma_0. \tag{6.11}$$

Таким образом, при циклотронном резонансе электропроводность в два раза меньше, чем в постоянном электрическом поле. Если бы высокочастотное электромагнитное поле было поляризовано по кругу, а не линейно, как у нас ($E_y = 0$), то $\sigma_{\mathrm{Re}}/\sigma_0 = 1$.

3. Рассмотрим теорию циклотронного резонанса в случае, когда энергетический спектр электронов состоит из ряда эквивалентных минимумов в $\boldsymbol{k}$-пространстве, вблизи которых поверхности постоянной энергии — эллипсоиды (электронные германий и кремний); в этом случае энергия электрона

$$\varepsilon(\boldsymbol{k}) = \frac{\hbar^2}{2}\left(\frac{k_1^2}{m_1} + \frac{k_2^2}{m_2} + \frac{k_3^2}{m_3}\right), \tag{6.12}$$

где волновый вектор $\boldsymbol{k}$ отсчитывается от точки минимума $\boldsymbol{k}_0$ и индексы 1, 2, 3 соответствуют прямоугольным осям, совпадающим с главными осями эллипсоида энергии. В случае электронного германия и кремния поверхности постоянной энергии (вблизи минимумов) — эллипсоиды вращения, так что $m_1 = m_2 = m_\perp$ и $m_3 = m_\parallel$, где $m_\perp$ и $m_\parallel$ — поперечная и продольная массы тензора эффективных масс. Используя (6.12), получим для составляющих скорости выражение

$$v_i = \frac{1}{\hbar}\frac{\partial\varepsilon}{\partial k_i} = \frac{\hbar k_i}{m_i} = \frac{p_i}{m_i} \quad (i = 1, 2, 3), \tag{6.13}$$

где квазиимпульс $\boldsymbol{p} = \hbar\boldsymbol{k}$.

Так как высокочастотное поле и сила трения $\boldsymbol{F}_R$ являются малыми возмущениями, достаточно рассмотреть квазиклассические уравнения движения в магнитном поле $\boldsymbol{H}$:

$$\frac{d(\hbar\boldsymbol{k})}{dt} = -\frac{e}{c}[\boldsymbol{v}\boldsymbol{H}]. \tag{6.14}$$

Используя (6.13), получим

$$\begin{aligned} m_1\frac{dv_1}{dt} &= -\frac{e}{c}v_2H_3+\frac{e}{c}v_3H_2,\\ m_2\frac{dv_2}{dt} &= -\frac{e}{c}v_3H_1+\frac{e}{c}v_1H_3,\\ m_3\frac{dv_3}{dt} &= -\frac{e}{c}v_1H_2+\frac{e}{c}v^2H_1, \end{aligned} \tag{6.14a}$$

где $H_1=H\alpha_1$, $H_2=H\alpha_2$, $H_3=H\alpha_3$ — проекции напряженности магнитного поля на координатные оси, а α_1, α_2 и α_3 — соответствующие направляющие косинусы.

Считая, что скорость электрона $\boldsymbol{v}\sim\exp(-\omega t)$, т. е., полагая

$$v_1=v_{10}e^{-i\omega t},\quad v_2=v_{20}e^{-i\omega t},\quad v_3=v_{30}e^{-i\omega t}, \tag{6.15}$$

получим, подставляя (6.15) в (6.14а):

$$\begin{aligned} -icm_1\omega v_{10}+eH_3v_{20}-eH_2v_{30}&=0,\\ eH_3v_{10}+icm_2\omega v_{20}-eH_1v_{30}&=0,\\ eH_2v_{10}-eH_1v_{20}-icm_3\omega v_{30}&=0. \end{aligned} \tag{6.16}$$

Эта система алгебраических линейных однородных уравнений относительно неизвестных амплитуд v_{10}, v_{20}, v_{30} имеет решение, отличное от нулевого только тогда, когда определитель системы (6.16) равен нулю [1]). Приравнивая определитель системы (6.16) нулю, получим характеристическое уравнение 3-й степени для ω, корни которого равны

$$\omega=0 \tag{6.17}$$

и

$$\omega^2=\omega_1^2\alpha_1^2+\omega_2^2\alpha_2^2+\omega_3^2\alpha_3^2\equiv\omega_c^2, \tag{6.17а}$$

где

$$\omega_1=\frac{eH}{c\sqrt{m_2m_3}},\quad \omega_2=\frac{eH}{c\sqrt{m_1m_3}},\quad \omega_3=\frac{eH}{c\sqrt{m_1m_2}}. \tag{6.17б}$$

Частота (6.17) $\omega=0$ соответствует движению электрона вдоль магнитного поля; частота (6.17а) $\omega=\omega_c$ — циклотронная частота, соответствующая движению электрона в плоскости, перпендикулярной к магнитному полю. Если магнитное поле совпадает, например, с первой осью, то $\alpha_1=1$, $\alpha_2=\alpha_3=0$ и $\omega_c=\omega_1$. Определяя циклотронную частоту (по максимуму поглощения в резонаторе) при разных ориентациях магнитного поля относительно кристалла, т. е. при разных α_1, α_2 и α_3, можно определить ω_1, ω_2 и ω_3, т. е. компоненты тензора эффективной массы m_1, m_2 и m_3.

Как отмечалось выше, поверхности постоянной энергии у электронного германия и кремния — эллипсоиды вращения, так что

[1]) Смирнов В. И., т. III, ч. 1, § 2, п. 10.

$m_1 = m_2 = m_\perp$ и $m_3 = m_\parallel$. Если магнитное поле $\boldsymbol{H}$ образует с осью 3 угол θ, то $\alpha_3^2 = \cos^2\theta$, $\alpha_1^2 + \alpha_2^2 = 1 - \alpha_3^2 = \sin^2\theta$ и из (6.17а) следует

$$\omega_c^2 = \left(\frac{eH}{c}\right)^2 \left[\frac{\cos^2\theta}{m_\perp^2} + \frac{\sin^2\theta}{m_\perp m_\parallel}\right]. \tag{6.18}$$

Измеряя ω_c для разных θ, можно определить $m_\perp$ и $m_\parallel$. Из-за того, что в $\boldsymbol{k}$-пространстве имеется несколько эквивалентных минимумов энергии, экспериментальные кривые поглощения имеют много пиков, соответствующих разным минимумам. Можно ввести циклотронную эффективную массу m_c^*, полагая

$$\omega_c = eH/m_c^* c; \tag{6.19}$$

тогда в случае (6.18)

$$\left(\frac{1}{m_c^*}\right)^2 = \frac{\cos^2\theta}{m_\perp^2} + \frac{\sin^2\theta}{m_\perp m_\parallel}, \tag{6.20}$$

т. е. m_c^* зависит не только от $m_\perp$ и $m_\parallel$, но и от угла θ, определяющего направление магнитного поля.

4. Выведем некоторое общее выражение для циклотронной эффективной массы m_c^*, справедливое при произвольном законе дисперсии для электронов (дырок) и позволяющее из опытов по циклотронному резонансу определять для них зависимость энергии ε от $\boldsymbol{k}$. В частности, формулы наши будут применимы к дырочным германию и кремнию, для которых $\varepsilon(\boldsymbol{k})$ имеет вид (IV.15.4), при котором, строго говоря, нельзя вводить тензор эффективной массы.

Рассмотрим движение электрона или дырки с произвольным законом дисперсии $\varepsilon(\boldsymbol{k})$ в однородном магнитном поле $\boldsymbol{H}$. Квазиклассические уравнения движения имеют вид

$$\frac{d}{dt}(\hbar\boldsymbol{k}) = \frac{e}{c}[\boldsymbol{v}\boldsymbol{H}], \tag{6.21}$$

$$\boldsymbol{v} = \frac{1}{\hbar}\nabla_{\boldsymbol{k}}\varepsilon(\boldsymbol{k}), \tag{6.21а}$$

где теперь e — алгебраическая величина заряда и $\nabla_{\boldsymbol{k}} \equiv \operatorname{grad}_{\boldsymbol{k}}$.

Если $\boldsymbol{H}$ направлено по оси z, то

$$\hbar\frac{dk_z}{dt} = \frac{e}{c}[\boldsymbol{v}\boldsymbol{H}]_z = 0$$

и, следовательно,

$$k_z = \text{const}. \tag{6.22}$$

С другой стороны, из (6.21) и (6.21а)

$$\frac{d\varepsilon(\boldsymbol{k})}{dt} = \nabla_{\boldsymbol{k}}\varepsilon\frac{d\boldsymbol{k}}{dt} = \boldsymbol{v}\frac{e}{c}[\boldsymbol{v}\boldsymbol{H}] = 0,$$

откуда

$$\varepsilon(\boldsymbol{k}) = \text{const}. \tag{6.23}$$

Таким образом, движение заряда в $\boldsymbol{k}$-пространстве происходит по кривой, определяемой пересечением поверхности постоянной энергии (6.23) с плоскостью (6.22).

Спроектируем обе части уравнения (6.21) в $\boldsymbol{k}$-пространстве на плоскость $k_z = \text{const}$:

$$\hbar \frac{dl}{dt} = \frac{e}{c} v_\perp H, \tag{6.24}$$

где dl — элемент «траектории», по которой происходит движение в $\boldsymbol{k}$-пространстве, а

$$v_\perp = \sqrt{v_x^2 + v_y^2} \tag{6.24а}$$

— проекция скорости на плоскость, перпендикулярную к магнитному полю (т. е. к оси z).

Если «траектория» в $\boldsymbol{k}$-пространстве замкнута, то период обращения заряда

$$T_c = \oint dt = \frac{\hbar c}{eH} \oint \frac{dl}{v_\perp}, \tag{6.25}$$

если подставить dt из (6.24).

С другой стороны, площадь сечения, ограниченная замкнутой траекторией,

$$S = \iint dk_x\, dk_y. \tag{6.26}$$

Перейдем в двойном интеграле от дифференциалов $dk_x\, dk_y$ к dl и $dk_\perp$; $k_\perp$ — элемент нормали к dl в плоскости $k_z = \text{const}$. Из общей формулы (6.21а) имеем

$$v_\perp = \frac{1}{\hbar} \frac{\partial \varepsilon}{\partial k_\perp}, \tag{6.27}$$

откуда

$$dk_\perp = \frac{1}{\hbar} \frac{d\varepsilon}{v_\perp}. \tag{6.27а}$$

Таким образом,

$$S = \iint dl\, dk_\perp = \frac{1}{\hbar} \int d\varepsilon \oint \frac{dl}{v_\perp} \tag{6.28}$$

или

$$\hbar \frac{\partial S}{\partial \varepsilon} = \oint \frac{dl}{v_\perp}. \tag{6.28а}$$

Подставляя это в соотношение (6.25), получим:

$$T_c = \frac{\hbar^2 c}{eH} \frac{\partial S}{\partial \varepsilon}. \tag{6.29}$$

Так как циклотронная частота $\omega_c = 2\pi/T_c$, то циклотронная эффективная масса (6.19) равна

$$m_c^* = \frac{\hbar^2}{2\pi} \frac{\partial S}{\partial \varepsilon}. \tag{6.30}$$

Измеряя m_c^* при различных ориентациях магнитного поля относительно кристалла, можно в принципе восстановить форму поверхности $\varepsilon(\boldsymbol{k})=\text{const}$. Если закон дисперсии $\varepsilon(\boldsymbol{k})$ задан в аналитической форме, но содержит ряд неизвестных параметров, как, например, в случаях (6.12) или (IV.15.4), то выражение (6.30) может быть использовано для экспериментального определения этих параметров. Для примера рассмотрим простой случай, когда поверхность постоянной энергии — сфера, так что

$$\varepsilon=\hbar^2k^2/2m^*, \tag{6.31}$$

где m^* — скалярная эффективная масса. В этом случае

$$S=\pi k_\perp^2=\pi\,(k^2-k_z^2)=\pi\left(\frac{2m^*\varepsilon}{\hbar^2}-k_z^2\right). \tag{6.31a}$$

Подставляя в (6.30), получим

$$m_c^*=m^*, \tag{6.32}$$

что соответствует (6.1а). Также просто из (6.30) получить выражение (6.20). В случае сложной энергетической поверхности $\varepsilon(\boldsymbol{k})=\text{const}$ (не сферы и не эллипсоида), производная $\partial S/\partial\varepsilon$ зависит, вообще говоря, от k_z, так что $m_c^*(k_z)$. Так как в кристалле имеются электроны с разными k_z, то максимуму кривой циклотронного поглощения соответствует $m_c^*(\bar{k}_z)$, где $\bar{k}_z\approx\bar{k}_x\approx\bar{k}_y$ по порядку величины определяются из условия $\varepsilon\approx k_0T$.

5. При достаточно низких температурах, удовлетворяющих неравенству (5.33), будут возбуждены магнитные осцилляторы только с малыми квантовыми числами N, поэтому квазиклассическое описание движения частицы (6.21) не может быть правильным. Полагая, как выше, $\omega_c=1{,}5\cdot10^{11}$ *рад/сек* и учитывая, что $2\mu^*H=\hbar\omega_c$, получим для (5.33)

$$T\leqslant1^\circ\,\text{K}. \tag{6.33}$$

Таким образом, при гелиевых температурах, при которых обычно ведутся опыты над циклотронным резонансом, мы находимся вблизи границы применимости квазиклассической теории.

С квантовой точки зрения циклотронный резонанс надо рассматривать как переходы электрона (дырки) с одной квантовой орбиты на другую под действием высокочастотного поля. Легко показать, что при этом возможны только переходы с изменением квантового числа магнитного осциллятора $\Delta N=\pm1$. С этой точки зрения явление естественно назвать *диамагнитным резонансом*.

В квантовой теории частота циклотронного резонанса $\omega_c^{(\text{кв})}$ определяется разностью энергий соседних термов $\mathscr{E}(N+1)-\mathscr{E}(N)\approx\hbar\omega_c^{(\text{кв})}$. В том случае, когда поверхности постоянной энергии — сферы, из (5.19) следует, что $\omega_c^{(\text{кв})}=2\mu^*H/\hbar=eH/m^*c=$

$= \omega_c$, т. е. не зависит от квантового числа N и совпадает с классической циклотронной частотой, определяемой из формулы (6.29). Точно так же обе эти частоты совпадают, когда поверхности постоянной энергии — эллипсоиды. Таким образом, в этих частных случаях формула (6.29) справедлива при любых сколь угодно низких температурах.

При более сложных энергетических поверхностях, например в p-Ge, разность $\mathcal{E}(N+1) - \mathcal{E}(N)$ зависит от номера уровня N, при этом $\omega_c^{(\text{кв})}$ совпадает с ω_c, определяемой по формуле (6.29), только при больших значениях квантовых чисел N.

Квантовая теория диамагнитного (циклотронного) резонанса рассмотрена в работе Латтинжера [1]).

§ 7. Контакт полупроводника с металлом. Выпрямление

1. Рассмотрим примесный полупроводник электронного типа с энергией диссоциации доноров $\varepsilon_D \ll k_0 T$, когда они практически все ионизованы. Если полупроводник приведен в контакт с металлом, то в результате перераспределения свободных электронов полупроводника в приконтактном слое потенциал [2]) его поверхности станет равным потенциалу поверхности металла. Так как полупроводник и металл свободно обмениваются электронами и находятся в состоянии статистического равновесия, то уровни химического потенциала в них совпадают. Известно, что работа выхода электрона из металла или полупроводника определяется как разность между потенциальной энергией электронов в вакууме и их химическим потенциалом внутри тела. Если работа выхода из металла $w_м$ больше работы выхода из полупроводника $w_п$, то потенциал в приконтактном слое полупроводника возрастает и энергетические зоны будут искривлены так, как это показано на рис. VI.5. При этом потенциал на поверхности полупроводника (металла) будет равен $V_0 = w_м - w_п > 0$, если потенциальную энергию свободных электронов в глубине полупроводника положить равной нулю, т. е. отсчитывать все энергии от нижнего края зоны проводимости полупроводника. В рассматриваемом случае концентрация свободных электронов в объеме полупроводника будет выше, чем в приконтактной области, которую мы в этом случае будем называть *запорным слоем*. Если $w_м < w_п$, то потенциал $V_0 < 0$, т. е. ниже нижнего края зоны проводимости в полупроводнике, так что энергетические зоны вблизи поверхности полупроводника искривлены противоположно тому, как показано на рис. VI.5. Кон-

---

[1]) Luttinger T. M.—Phys. Rev, 1955, v. 102, p. 1030.

[2]) В дальнейшем мы для краткости называем потенциальную энергию электрона $V = -e\varphi$ (где φ — электростатический потенциал) потенциалом.

центрация электронов в объеме полупроводника будет ниже, чем в приконтактном слое, и мы этот слой будем называть *антизапорным*.

Очевидно, что в первом случае ($w_м > w_п$), представленном на рис. VI.5, электростатический потенциал на поверхности $\varphi_0 = -\frac{V_0}{e} < 0$; во втором случае ($w_м < w_п$) — $\varphi_0 > 0$. Так же

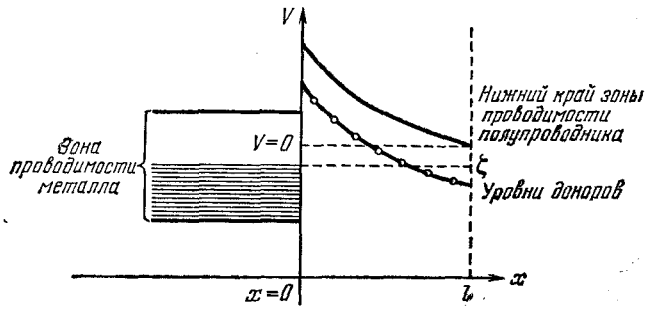


Рис. VI.5.

понятно, что первый случай ведет в случае дырочного полупроводника к образованию антизапорного, второй — запорного слоя.

Так как последовательное включение в цепь малого сопротивления (антизапорного слоя) никаких особенностей не вызывает, мы будем в дальнейшем интересоваться только запорными слоями, которые, как мы увидим, обладают выпрямляющими свойствами.

Потенциал запорного слоя V удовлетворяет уравнению Пуассона:

$$\nabla^2 V = \frac{4\pi}{\varepsilon} e\rho, \tag{7.1}$$

где ε — диэлектрическая постоянная, ρ — плотность электрического заряда. Предполагая, что все величины зависят только от координаты x, перпендикулярной к поверхности полупроводника, имеем

$$\frac{d^2V}{dx^2} = \frac{4\pi e^2}{\varepsilon}[n_D - n(x)], \tag{7.2}$$

где n_D и $n(x)$ — концентрации положительно заряженных (ионизованных) доноров и свободных электронов. Из соображений нейтральности следует, что в объеме полупроводника $n(x) = n = n_D$. В запорном слое концентрация электронов, согласно закону Больцмана, равна

$$n(x) = n \exp\left(-\frac{V(x)}{k_0 T}\right). \tag{7.3}$$

Вводя безразмерный потенциал

$$\Phi(x)=\frac{V(x)}{k_0T}. \tag{7.4}$$

и характеристическую *дебаевскую длину* (*радиус*)

$$l_D=\sqrt{\frac{\varepsilon k_0T}{4\pi e^2n}}, \tag{7.4а}$$

запишем уравнение (7.2) в виде

$$\frac{d^2\Phi}{dx^2}=\frac{1}{l_D^2}\left[1-e^{-\Phi(x)}\right]. \tag{7.4б}$$

Последнее уравнение содержит единственный параметр l_D; поэтому из соображений размерности следует, что толщина запорного слоя порядка l_D. Дебаевская длина l_D играет очевидную роль *радиуса экранирования потенциала* (заряда).

Если, например, $\Phi(x)\ll 1$, т. е. $V(x)\ll k_0T$ (редко реализующийся на практике случай), то, разлагая в (7.4б) экспоненту в ряд, получим

$$\frac{d^2V}{dx^2}=\frac{1}{l_D^2}V. \tag{7.5}$$

Учитывая, что $V(0)=V_0$ и $V(\infty)=0$, получим

$$V(x)=V_0e^{-x/l_D}, \tag{7.6}$$

т. е. l_D—длина, на которой потенциал (заряд) уменьшается в e раз.

Для $T=300°\mathrm{K}$, $n=10^{16}$ *см*$^{-3}$ и $\varepsilon=16$ (Ge) дебаевская длина $l_D=4\cdot10^{-6}$ *см*, т. е. во много раз больше постоянной решетки. Для нормальных металлов l_D порядка постоянной решетки, так что, строго говоря, вводить это понятие в данном случае нельзя.

2. При наличии электрического поля E_x и градиента концентрации свободных электронов $\frac{dn(x)}{dx}$ плотность электрического тока

$$j=\sigma E_x+(-e)\left(-D\frac{dn(x)}{dx}\right), \tag{7.7}$$

где D—коэффициент диффузии; удельная электропроводность

$$\sigma=en(x)\mu, \tag{7.7а}$$

где μ—подвижность электронов, (6.6б) с которой мы познакомимся подробнее в гл. IX.

В выражении (7.7) ток σE_x называется *полевым*, а ток $(-e)\left(-D\frac{dn}{dx}\right)$—*диффузионным*.

В статистическом равновесии полный ток всегда равен нулю. Это значит, что, например, в запорном слое полевой ток урав-

новешивается диффузионным [1]). В этом случае из (7.7) и (7.7а) получим

$$-en(x)\mu\frac{d\varphi}{dx}+eD\frac{dn(x)}{dx}=0, \tag{7.8}$$

где E_x заменено на $-\frac{d\varphi}{dx}$. Интегрируя (7.8), получим

$$n(x)=ne^{\mu\varphi/D}, \tag{7.9}$$

так как $n(x)=n$ при $\varphi=0$. С другой стороны, в невырожденном электронном газе, по Больцману,

$$n(x)=n_0\exp\left(-\frac{V}{k_0T}\right)=n\exp\frac{e\varphi}{k_0T}. \tag{7.10}$$

Из сравнения (7.9) с (7.10) следует *соотношение Эйнштейна* — связь между подвижностью и коэффициентом диффузии

$$\mu=\frac{e}{k_0T}D. \tag{7.11}$$

3. Рассмотрим теперь явления, возникающие при прохождении тока через запорный слой. Электроны, движущиеся в полупроводнике, испытывают столкновения с фотонами (колебаниями решетки), примесными центрами и другими нарушениями периодической структуры кристаллической решетки. В результате таких столкновений электроны резко меняют направление своего движения. Можно ввести понятие о средней длине свободного пробега электрона проводимости l, понимая под этим его средний путь между последовательными столкновениями [2]). Очевидно, $l=v\tau$, где v — скорость и τ — среднее время свободного пробега электрона (§ 6, п. 1).

Рассматривая прохождение тока через запорный слой, следует различать два случая: первый — когда длина свободного пробега электрона много больше толщины запорного слоя, и второй — когда выполняется обратное соотношение. Ввиду большого сопротивления запорного слоя, обедненного носителями тока, вся приложенная в цепи разность потенциалов U практически приходится на него. В первом случае (диодная теория выпрямления, Бете, 1942), когда длина свободного пробега электронов велика, можно очень просто вычислить поток электронов, падающий из объема полупроводника на его границу ($x=0$). Число электронов в объеме полупроводника со скоростями, лежащими между v_x и v_x+dv_x, v_y и v_y+dv_y, v_z и v_z+dv_z, которые проходят в 1 *сек* сквозь 1 *см*^{2} поверхности, перпендику-

---

[1]) Этот случай более подробно рассматривается в следующем пункте.

[2]) В следующей главе понятие средней длины свободного пробега будет строго определено из кинетического уравнения.

лярной к оси x, равно

$$dn = n\left(\frac{m^*}{2\pi k_0 T}\right)^{3/2} e^{-\frac{m^*\left(v_x^2+v_y^2+v_z^2\right)}{2k_0T}} v_x dv_x dv_y dv_z, \qquad (7.12)$$

где n — концентрация электронов в глубине полупроводника [1]). Для того чтобы определить поток электронов, достигающих поверхности полупроводника (при $x=0$), надо проинтегрировать выражение (7.12) по v_y и v_z от $-\infty$ до $+\infty$ и по v_x от значения, определяемого из равенства

$$m^* v_x^2/2 = V_0 - eU = -e(\varphi_0 + U) > 0, \qquad (7.13)$$

до $v_x = \infty$.

Здесь U есть изменение отрицательного электростатического потенциала φ_0 на границе $x=0$ за счет приложенной к запорному слою разности потенциалов. В самом деле, если $|v_x|$ имеет меньшее значение, чем определяемое из (7.13), то такой электрон не сможет преодолеть потенциальный барьер запорного слоя и достичь поверхности полупроводника.

Таким образом, поток электронов из полупроводника в металл

$$Q = n\left(\frac{m^*}{2\pi k_0 T}\right)^{3/2} \int\limits_{\sqrt{2(V_0-eU)/m^*}}^{\infty} v_x \exp\left(-\frac{m^* v_x^2}{2k_0T}\right) dv_x \times$$

$$\times \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \exp\left(-\frac{m^*\left(v_y^2+v_z^2\right)}{2k_0T}\right) dv_y\, dv_z = \frac{1}{4} n v_T \exp\left(-\frac{V_0}{k_0T}\right) \exp\frac{eU}{k_0T} =$$

$$= Q_0 \exp\frac{eU}{k_0T}, \qquad (7.14)$$

где $v_T = \left(\frac{8k_0T}{\pi m^*}\right)^{1/2}$ — средняя тепловая скорость электронов, nv_T — плотность потока электронов в глубине полупроводника, $Q_0 = \left(\frac{1}{4}\right) nv \exp\left(-\frac{V_0}{k_0T}\right)$ — плотность потока электронов из полупроводника в металл через границу $x=0$ при $U=0$.

Интегралы по v_y и v_z вычисляются по формуле (II.6.1); интеграл по v_x вычисляется элементарно, если посредством преобразования $v_x dv_x = {}^1/_2 dv_x^2$ перейти к переменной интегрирования $t = v_x^2$. Так как в равновесном состоянии результирующий поток электронов через границу $x=0$ равен нулю, а приложенный к запорному слою потенциал U не влияет на состояние электронов в металле, то поток электронов из металла в полупроводник всегда равен Q_0.

Если $U > 0$, т. е. металл служит положительным электродом, а полупроводник отрицательным, то барьер V_0 уменьшается, так

---

[1]) Ансельм А. И., с. 63.

что $Q > Q_0$. Результирующий поток электронов течет из полупроводника в металл, а электрический ток — в противоположном направлении. Если мы условимся считать этот ток положительным, то он равен

$$j = e(Q - Q_0) = j_0\left(\exp\frac{eU}{k_0T} - 1\right), \tag{7.15}$$

где $j_0 = eQ_0$.

Это направление приложенного потенциала называют *пропускным*, а ток — *прямым* или *пропускным*. При $U < 0$, когда металл является отрицательным электродом, потенциальный барьер на границе полупроводника увеличивается и результирующий поток электронов направлен из металла в полупроводник. В этом случае электрический ток определяется тем же выражением (7.15), но имеет отрицательную величину.

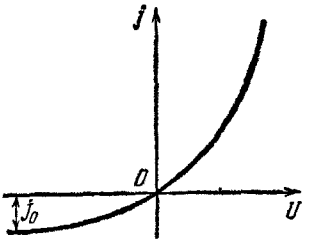


Рис. VI. 6.

В этом случае мы говорим о *запорном направлении*, а ток называется *обратным* или *запорным*.

На рис. VI.6 схематически представлена, согласно (7.15), вольтамперная характеристика диодного выпрямителя.

При $U < 0$ в запорном направлении при возрастании $|U|$ ток стремится к насыщению $(-j_0)$. При положительных $U \gg \frac{k_0T}{e}$ ток возрастает по экспоненциальному закону, однако только до значений $U < |\varphi_0|$, при которых еще существует запорный слой. При U, близких к $|\varphi_0|$, сопротивление запорного слоя настолько уменьшается, что значительная часть приложенного напряжения будет приходиться на объем полупроводника. Выражение (7.15), не учитывающее этого, при таких напряжениях уже неприменимо.

Рассмотрим теперь так называемую *диффузионную теорию выпрямления* (Б. И. Давыдов, Шоттки, С. И. Пекар, 1939), применимую тогда, когда длина свободного пробега электрона l много меньше толщины запорного слоя l_D. В этом случае ток определяется выражением (7.7). Кроме того, мы должны считать, что на поверхности

$$\varphi(0) = \varphi_0 + U. \tag{7.16}$$

Из (7.7) имеем

$$\frac{dn(x)}{dx} - \frac{e}{k_0T}\frac{d\varphi}{dx}n(x) - \frac{j}{\mu k_0T} = 0, \tag{7.17}$$

где использовано соотношение Эйнштейна (7.11).

Распределение электростатического потенциала $\varphi(x)$ связано, конечно, с распределением концентрации $n(x)$. Однако формально

мы можем рассматривать (7.17) как линейное дифференциальное уравнение первого порядка для неизвестной функции $n(x)$, считая $\frac{d\varphi}{dx}$ заданной функцией x. Решение (7.17) имеет вид [1])

$$n(x) = n \exp \frac{e\varphi(x)}{k_0 T} - \frac{j}{\mu k_0 T} \int_x^\infty \exp \frac{e}{k_0 T} [\varphi(x) - \varphi(\xi)]\, d\xi, \qquad (7.18)$$

где единственная константа интегрирования выбрана так, чтобы удовлетворялось граничное условие $n(\infty) = n$ постоянной концентрации в глубине полупроводника. Решение (7.18) может быть проверено прямой подстановкой в уравнение (7.17). При дифференцировании по x интеграла в (7.18) необходимо учесть зависимость от x как подынтегральной функции, так и нижнего предела интеграла [2]).

Применим теперь решение (7.18) к поверхности полупроводника, т. е. при $x = 0$; тогда

$$n_0 \equiv n(0) = n \exp \frac{e\varphi_0}{k_0 T} \exp \frac{eU}{k_0 T} - \frac{j}{\mu k_0 T} \int_0^\infty \exp \frac{e}{k_0 T} [\varphi(0) - \varphi(\xi)]\, d\xi, \qquad (7.19)$$

где использовано выражение (7.16).

Так как в рассматриваемом нами случае потоки электронов из металла в полупроводник и обратный (каждый в отдельности) много больше результирующего потока $(1/e)\, j$, то на границе $x = 0$ практически сохраняется равновесная концентрация электронов, т. е.

$$n_0 = n \exp \frac{e\varphi_0}{k_0 T}. \qquad (7.20)$$

Решая (7.19) относительно j и используя (7.20), получим

$$j = j_0 \left[\exp \frac{eU}{k_0 T} - 1 \right], \qquad (7.21)$$

где

$$j_0 = \frac{n_0 \mu k_0 T}{\int_0^\infty \exp \frac{e}{k_0 T} [\varphi(0) - \varphi(\xi)]\, d\xi}. \qquad (7.21a)$$

Как мы увидим ниже, интеграл, определяющий j_0, слабо зависит от U, поэтому соотношение между j и U в диффузионной теории выпрямления (7.21) имеет тот же вид, что и в диодной (7.15). В запорном слое электростатический потенциал $\varphi(x)$ отри-

[1]) Смирнов В. И., т. II, § 1, п. 6.
[2]) Там же, § 8, п. 83.

цателен и имеет наибольшее абсолютное значение при $x=0$; поэтому показатель экспоненты в интеграле (7.21а) отрицателен и при возрастании ξ быстро возрастает по абсолютной величине. В силу этого при вычислении интеграла можно разложить $\varphi(\xi)$ в ряд по ξ и ограничиться линейным приближением:

$$\varphi(\xi)=\varphi(0)+\left(\frac{d\varphi}{d\xi}\right)_0\xi=\varphi(0)+\left(\frac{d\varphi}{dx}\right)_0\xi. \tag{7.22}$$

Тогда

$$\int_0^\infty \exp\left\{\frac{e}{k_0T}[\varphi(0)-\varphi(\xi)]\right\}d\xi\approx$$

$$\approx\int_0^\infty \exp\left[-\frac{e}{k_0T}\left(\frac{d\varphi}{dx}\right)_0\xi\right]d\xi=\frac{k_0T}{e}\frac{1}{\left(\frac{d\varphi}{dx}\right)_0}. \tag{7.23}$$

Подставляя (7.23) в (7.21а), получим

$$j_0=en_0\mu\left(\frac{d\varphi}{dx}\right)_0 \tag{7.24}$$

— полевой ток на границе полупроводника. В (7.21) можно не учитывать слабую зависимость $\left(\frac{d\varphi}{dx}\right)_0$ от U по сравнению с экспоненциальным членом, тогда выражение (7.21) аналогично (7.15).

Таким образом, и в случае диффузионной теории выпрямления вольтамперная характеристика имеет вид, изображенный на рис. VI.6 однако при той же высоте барьера φ_0 ток j_0 значительно меньше.

Образование запирающего слоя вблизи поверхности полупроводника не всегда связано с разностью работ выхода металла и полупроводника. Нередки случаи, когда запорный слой образуется в результате заряжения поверхности полупроводника электронами, локализованными на поверхностных уровнях (гл. V, § 2, п. 3). В этом случае поверхностный барьер V_0 не зависит от природы металлического контакта (Бардин, 1947).

В ряде технически важных случаев запорный слой химического происхождения (пленки окислов, шеллак, лак и др.). Теория выпрямления для этого случая (Мотт) сходна с изложенной выше.

§ 8. Свойства p—n-переходов

1. В предыдущем параграфе мы исследовали электрические свойства контакта полупроводника с металлом. Представляется естественным рассмотреть теперь контакт между двумя полупроводниками, например с разным типом проводимости — электронным и дырочным. В наиболее чистом виде контакт полупроводников n- и p-типа может быть осуществлен, если внутри одного и

того же вещества (кристалла) создать две граничащие друг с другом области: одну—с электронной, другую—с дырочной проводимостью. Такая граница p и n областей внутри полупроводника называется *p—n-переходом* (или *электронно-дырочным переходом*).

В большинстве полупроводниковых радиотехнических приборов используются p—n-переходы в германии или кремнии; поэтому технология изготовления p—n-переходов в них, на которой мы останавливаться не будем, достигла высокого совершенства.

Рассмотрим идеализированный p—n-переход в Ge: слева от плоскости $x=0$ концентрация акцепторов (например, атомов индия) постоянна и равна n_A, справа от $x=0$ имеется постоянная концентрация доноров (например, атомов сурьмы), равная n_D. При не слишком низких температурах все рассматриваемые примеси ионизованы, так что в глубине левой p-области концентрация дырок $p_p \approx n_A$, а в глубине правой n-области концентрация электронов проводимости $n_n \approx n_D$. Подвижные дырки будут диффундировать из левой p-области в правую n-область; наоборот, электроны проводимости будут переходить сквозь плоскость $x=0$ справа налево. В результате слева от $x=0$ образуется диффузный слой отрицательного заряда, справа от $x=0$—слой положительного заряда. Образовавшийся двойной слой создает скачок потенциала, препятствующий дальнейшему переходу дырок слева направо и электронов справа налево.

Наряду с примесной проводимостью в германии имеется и некоторая собственная проводимость. Пусть n_p и p_n—соответственно равновесные концентрации электронов в глубине p-области и дырок в глубине n-области. В любой точке рассматриваемого нами p—n-перехода концентрация электронов n и дырок p удовлетворяет соотношению (2.20):

$$np = n_i^2, \tag{8.1}$$

где n_i—концентрация носителей тока собственной проводимости. Из (8.1)

$$\lg p - \lg n_i = \lg n_i - \lg n. \tag{8.1а}$$

На рис. VI.7 представлены в логарифмическом масштабе концентрации электронов и дырок в p—n-переходе. Мы видим, что кривые $\lg p(x)$ и $\lg n(x)$ симметричны относительно прямой $\lg n_i$, что непосредственно следует из (8.1а). Плоскости x_p и x_n определяют границы двойного слоя или быстрого изменения потенциала в p—n-переходе. Во всех практических случаях $n_p \ll p_p$ и $p_n \ll n_n$[1]). Если, например[2]), $n_A = p_p = 10^{16}$ и $n_D = n_n = 10^{14}$,

[1]) Поэтому мы выше положили $p_p = n_A$ и $n_n = n_D$.

[2]) Здесь и дальше концентрации даются в $см^{-3}$.

то из (8.1)

$$n_p = \frac{n_i^2}{p_p} \approx 10^{10}, \quad p_n = \frac{n_i^2}{n_n} \approx 10^{12},$$

так как в германии при комнатной температуре $n_i \approx 10^{13}$. Для p—n-перехода, аналогично запорному слою, может быть написано соответствующее уравнение Пуассона (7.2), которое может быть решено в ряде предельных случаев. Обозначая опять электростатический потенциал через $\varphi(x)$ и полагая в глубине n-области $\varphi(+\infty) = 0$, получим по Больцману

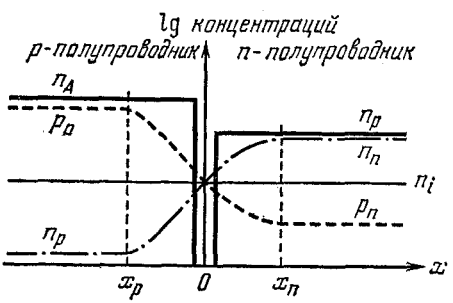


Рис. VI. 7.

$$p(x) = p_n \exp\left(-\frac{e\varphi(x)}{k_0 T}\right), \tag{8.2}$$

$$n(x) = n_n \exp\left(+\frac{e\varphi(x)}{k_0 T}\right). \tag{8.2a}$$

Если $\varphi(-\infty) \equiv \varphi_0$, то скачок потенциала на p—n-переходе будет равен

$$\varphi_0 = -\frac{k_0 T}{e} \lg \frac{p_p}{p_n} = \frac{k_0 T}{e} \lg \frac{n_p}{n_n}. \tag{8.3}$$

При комнатной температуре и приведенных выше концентрациях

$$\varphi_0 = -0{,}026 \cdot \lg 10^4 = -0{,}24 \text{ в}. \tag{8.3a}$$

Поскольку мы рассматривали p—n-переход в состоянии статистического равновесия, химический потенциал ζ должен

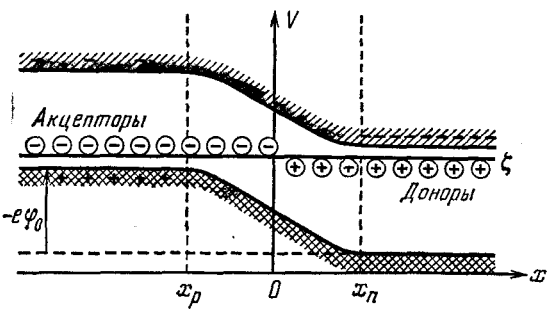


Рис. VI. 8.

проходить на одном уровне как в p, так и в n-области. На рис. VI.8 изображена потенциальная энергия электрона $V(x) = -e\varphi(x)$

вдоль $p-n$-перехода. Область существенного изменения потенциала (x_n-x_p) порядка дебаевской длины l_D.

2. Рассмотрим теперь прохождение тока через $p-n$-переход и покажем, что он обладает выпрямляющими свойствами[1]). Удерживая потенциал n-области равным нулю (например, заземлив ее), приложим к p-области положительный электростатический потенциал U. Так как сопротивление самого $p-n$-перехода относительно велико, будем считать, что вся разность потенциалов U падает на область (x_n-x_p). Положительный потенциал U уменьшает скачок отрицательного потенциала φ_0 (8.3) и вызывает дополнительный поток дырок из p-области, который, проникая в n-область, постепенно рекомбинирует с электронами, так что концентрация дырок падает до равновесного значения p_n. Аналогичная картина наблюдается для электронов, движущихся в n-области навстречу дыркам и проникающим в p-область. В результате чисто дырочный ток j_p в p-области сменяется чисто электронным током j_n в n-области. В стационарном случае сумма обоих токов j_p+j_n должна быть постоянна во всех точках $p-n$-перехода. Такое проникновение неосновных носителей тока в полупроводник (дырок в n-область и обратно) называется *инжекцией*. Работа большинства полупроводниковых приборов основана на явлении инжекции.

В рассматриваемом нами случае, когда дырки в p-области и электроны в n-области движутся навстречу друг другу, сопротивление $p-n$-перехода уменьшается аналогично тому, как это имело место при движении электронов из полупроводника в металл (§ 7). Таким образом, ток будет возрастать быстрее, чем по линейному закону Ома, т. е. мы имеем случай прямого направления $U>0$.

Мы рассмотрим прохождение тока в прямом направлении в случае малой скорости рекомбинации неосновных носителей, т. е. их глубокого проникновения в полупроводник (Шокли). Точнее, переходная область, в которой ток j_p сменяется током j_n, много больше «толщины» $p-n$-перехода x_n-x_p. Это значит, что в точках x_p и x_n составляющие токов $j_n(x_p)=j_n(x_n)$ и $j_p(x_p)=j_p(x_n)$. При $n_n=p_p$, $j_n=j_p$. При $U=0$ в состоянии равновесия

$$n(x_p)=n_p=n_n\exp\frac{e\varphi_0}{k_0T}\quad(\varphi_0<0),\tag{8.4}$$

так как плотность электронов подчиняется закону Больцмана (8.2а). При $U>0$ поток электронов из n- в p-область больше встречного потока электронов, но разность между этими потоками, т. е. результирующий поток, мала по сравнению с каждым из них; поэтому распределение плотности и в этом случае

[1]) Для простоты мы рассматриваем симметричный случай, когда $n_A=n_D$ и, следовательно, $x_n=|x_p|$.

приближенно описывается законом Больцмана:

$$n(x_p) = n_n \exp \frac{e\varphi_0 + eU}{k_0 T} = n_p \exp \frac{eU}{k_0 T}. \qquad (8.4a)$$

Таким образом, если напряжение $U > 0$, т. е. приложено в прямом направлении, то, согласно (8.4а), концентрация электронов на границе x_p возрастет и они будут диффундировать в глубь p-области, постепенно рекомбинируя с дырками.

Токи электронов в сечениях x_p и x_n, согласно (7.7), равны

$$j_n(x_p) = e\left[\mu_n n(x_p) E_x + D_n \frac{dn}{dx}\right], \qquad (8.5)$$

$$j_n(x_n) = e\left[\mu_n n(x_n) E_x + D_n \frac{dn}{dx}\right], \qquad (8.5a)$$

где все величины в правых частях (8.5) и (8.5а) относятся соответственно к сечениям x_p и x_n.

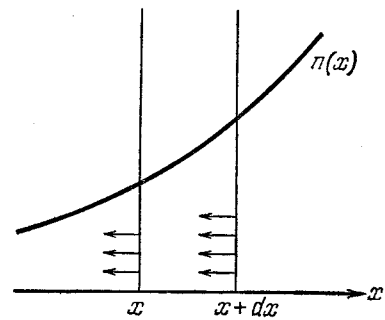


Рис. VI. 9.

В n-области, где концентрация электронов велика, даже слабое электрическое поле E_x вызовет большой омический ток, так что можно будет пренебречь диффузионным током и практически: $j_n(x_n) \approx e\mu_n n_n E_x$. С другой стороны, если $\exp\left[\frac{e(\varphi_0+V)}{k_0 T}\right] \ll 1$, то $\frac{n(x_p)}{n_n} \ll 1$ и, следовательно, омический ток в $j_n(x_p)$ мал по сравнению с $e\mu_n n_n E_x = j_n(x_n)$. Так, $j_n(x_p) \approx j_n(x_n)$ и ток $j_n(x_p)$ практически весь диффузионный.

Вычислим диффузионный ток электронов в сечении x_p. Для этого мы рассмотрим вначале баланс электронов в слое dx на рис. VI.9. Поток электронов, втекающих в слой dx через сечение $x+dx$, равен $D_n \frac{dn(x+dx)}{dx}$, где D_n—коэффициент диффузии, а поток электронов, вытекающих через сечение x, равен $D_n \frac{dn(x)}{dx}$; поэтому в стационарном случае

$$D_n \frac{dn(x+dx)}{dx} - D_n \frac{dn(x)}{dx} = (\gamma np - g)\, dx, \qquad (*)$$

где g—число электронов (дырок), образующихся в 1 $см^3$ в 1 $сек$ в результате действия теплового возбуждения, а γnp—число электронов, исчезающих в 1 $см^3$ в 1 $сек$ в результате рекомбинации их с дырками (γ—коэффициент объемной рекомбинации); g одинаково во всех точках p—n-перехода; в глубине p-области генерация и рекомбинация полностью компенсируют друг друга;

поэтому

$$g = \gamma n_p p_p.$$

В $x \approx x_p$ концентрация дырок $p \approx p_p$. Так как

$$\frac{dn(x+dx)}{dx} = \frac{dn(x)}{dx} + \frac{d^2n(x)}{dx^2}dx,$$

то из (*) следует

$$D_n \frac{d^2n'(x)}{dx^2} = \gamma p_p [n(x) - n_p]$$

или

$$\frac{d^2n(x)}{dx^2} = \frac{1}{\mathscr{L}_n^2}[n(x) - n_p], \tag{8.6}$$

где

$$\mathscr{L}_n = \sqrt{\frac{D_n}{\gamma p_p}} \tag{8.6а}$$

— *диффузионная длина для электронов* в p-области.

В нестационарном случае в отсутствие тока скорость исчезновения электронов $-\left(\frac{\partial n}{\partial t}\right)$ пропорциональна отклонению концентрации n от ее равновесного значения[1])

$$-\frac{\partial n}{\partial t} = \frac{n - n_p}{\tau_n}, \tag{8.7}$$

где τ_n — *время жизни электронов* в p-области.

С другой стороны,

$$-\frac{\partial n}{\partial t} = \gamma n p_n - g = \gamma p_p (n - n_p). \tag{8.8}$$

Из сравнения (8.7) с (8.8) следует, что

$$\tau_n = 1/\gamma p_p, \tag{8.9}$$

но тогда

$$\mathscr{L}_n = \sqrt{D_n \tau_n}. \tag{8.10}$$

Теперь мы можем точнее сформулировать используемое нами условие малой рекомбинации или глубокой инжекции:

$$\mathscr{L}_n \gg |x_p|. \tag{8.11}$$

Решение уравнения (8.6), удовлетворяющее граничному условию $n(-\infty) = n_p$, имеет вид

$$n(x) = n_p + C \exp \frac{x}{\mathscr{L}_n}, \tag{8.12}$$

где C — константа интегрирования.

Поток электронов в сечении x_p справо налево равен

$$D_n \frac{dn(x)}{dx}\Big|_{x=x_p} = \frac{D_n}{\mathscr{L}_n} C e^{x/\mathscr{L}_n} = \frac{D_n}{\mathscr{L}_n}[n(x_p) - n_p], \tag{8.13}$$

---

[1]) Во всяком случае, при малых отступлениях концентраций от равновесного значения n_p.

поэтому ток

$$j_n(x_p) = \frac{eD_n}{\mathscr{L}_n}[n(x_p) - n_p]. \qquad (8.14)$$

В статистическом равновесии $n(x_p) = n_p$ и ток $j_n = 0$. Если повысить электростатический потенциал в точке x_p на U, то потенциальная энергия электрона понизится на $-eU$ и концентрация электронов повысится в соответствии с (8.4а). Из (8.14) и (8.4а) следует

$$j_n(x_p) \approx \frac{eD_n n_p}{\mathscr{L}_n}\left[\exp\frac{eU}{k_0 T} - 1\right]. \qquad (8.15)$$

Рассуждая аналогично, получим

$$j_p(x_n) \approx \frac{eD_p p_n}{\mathscr{L}_p}\left[\exp\frac{eU}{k_0 T} - 1\right]. \qquad (8.16)$$

Как мы отмечали, $j_p(x_n) \approx j_p(x_p)$, поэтому полный ток в прямом направлении

$$j = j_n(x_p) + j_p(x_p) = j_0\left[e^{\frac{eU}{k_0 T}} - 1\right], \qquad (8.17)$$

где

$$j_0 \equiv \frac{eD_n n_p}{\mathscr{L}_n} + \frac{eD_p p_n}{\mathscr{L}_p}. \qquad (8.17а)$$

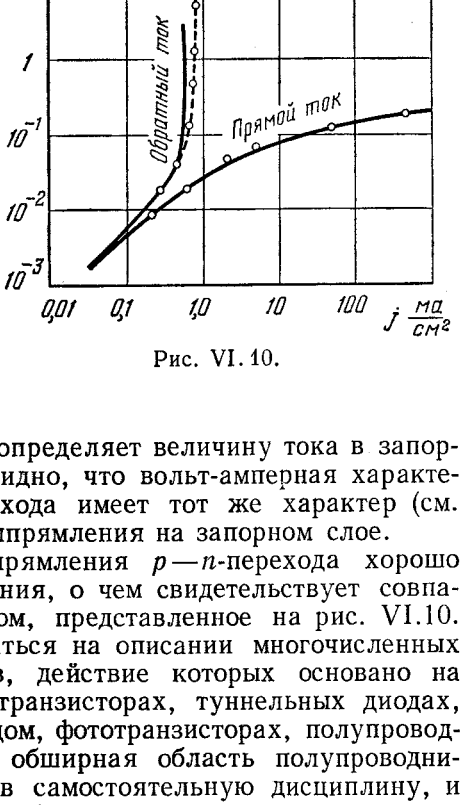


Рис. VI.10.

При $U < 0$ выражение (8.17) определяет величину тока в запорном направлении. Из (8.17) видно, что вольт-амперная характеристика в случае p—n-перехода имеет тот же характер (см. рис. VI.6), что и в случае выпрямления на запорном слое.

Предпосылки теории выпрямления p—n-перехода хорошо выполняются в случае германия, о чем свидетельствует совпадение теории с экспериментом, представленное на рис. VI.10.

Мы не можем останавливаться на описании многочисленных полупроводниковых приборов, действие которых основано на применении p—n-перехода: транзисторах, туннельных диодах, фотоэлементах с p—n-переходом, фототранзисторах, полупроводниковых лазерах и т. д. Эта обширная область полупроводниковой электроники выросла в самостоятельную дисциплину, и для ее изучения необходимо обратиться к специальным сочинениям[1]).

[1]) Хорошее изложение этих вопросов можно найти в книге Пикус Г. Е. Основы теории полупроводниковых приборов.— М.: Наука, 1965.

§ 9. Генерация и рекомбинация носителей тока. Квазиуровни Ферми

1. При рассмотрении статистически равновесных свойств полупроводника (§ 2) было отмечено, что они не зависят от механизма взаимодействия носителей тока друг с другом и с решеткой. При исследовании неравновесных свойств, в частности явлений переноса свободных зарядов, механизм взаимодействия становится существенным.

В собственных полупроводниках, лишенных примесных центров, свободные электроны и дырки возникают в результате перехода электронов из валентной зоны в зону проводимости. Такая *генерация электронно-дырочных пар* может быть связана с тепловым возбуждением электрона в валентной зоне или поглощением им электромагнитной энергии от внешнего источника, или передачей ему энергии от электрона проводимости в сильном электрическом поле и т. п. Исчезновение таких электронно-дырочных пар происходит в результате *рекомбинации свободного электрона с дыркой*. Такие процессы *прямой рекомбинации* электрона и дырки непосредственно через запрещенную зону затруднены тем, что при этом должна выделиться сравнительно большая энергия порядка ширины запрещенной зоны ε_G. Эта энергия может быть излучена в виде фотона частоты $\omega_s \approx \varepsilon_G/\hbar$ (*излучательный переход*) или, что много менее вероятно, — в виде множества фононов с энергиями порядка $\hbar\omega_q \approx k_0 T \ll \varepsilon_G$ (*безызлучательный* переход).

Процессы прямой рекомбинации могут также идти с участием третьей свободной частицы (электрона или дырки). При таком тройном столкновении энергия ε_G, выделяемая при рекомбинации, передается в виде кинетической энергии третьей частице. Это явление аналогично оже-процессу в изолированном атоме [1]) (Пьер Оже), поэтому оно получило название оже-рекомбинации.

Возможна, наконец, рекомбинация, связанная с одновременным участием фотона и фонона (последний практически меняет только импульс свободной частицы). Такие непрямые переходы будут нами подробнее рассмотрены при изучении поглощения света полупроводниками (гл. VII, § 3). Очевидно, что изменение числа свободных электронов (дырок) в 1 $см^3$ в 1 *сек* в результате прямой рекомбинации (без участия третьей частицы) пропорционально числу столкновений электронов с дырками, которое пропорционально их концентрациям n и p. Таким образом,

$$-\left(\frac{dn}{dt}\right)_r = -\left(\frac{dp}{dt}\right)_r = \gamma np - g \qquad (9.1)$$

[1]) Шпольский Э. В. Атомная физика.— 4 изд.— М.: Наука, 1974, т. II, с. 299.

Здесь γ — коэффициент объемной рекомбинации, имеющий размерность $см^3 сек^{-1}$, g — скорость генерации электронно-дырочных пар в 1 $см^3$ в результате теплового возбуждения электронов из валентной зоны [1]). В состоянии статистического равновесия полупроводника $n = n_0$, $p = p_0$ и $\gamma n_0 p_0 - g = 0$, т. е. скорость тепловой генерации

$$g = \gamma n_0 p_0, \tag{9.2}$$

что, конечно, имеет место и в неравновесном случае. Таким образом, (9.1) приобретает вид

$$-\left(\frac{dn}{dt}\right)_r = \gamma (np - n_0 p_0). \tag{9.3}$$

В гл. V, § 2 упоминалось о мелких и глубоких примесных уровнях в германии и кремнии. Глубокие примесные уровни могут существенно облегчить процесс рекомбинации свободного электрона с дыркой, так как при переходе электрона (дырки) на глубокий примесный уровень должна выделиться меньшая энергия, чем при прямой рекомбинации зона $\rightleftarrows$ зона. Такую *непрямую* (*ступенчатую*) рекомбинацию можно себе представить, например, так: свободный электрон захватывается глубоким донорным уровнем, после чего нейтральный донор захватывает дырку, так что в итоге происходит рекомбинация свободного электрона с дыркой. Следует отметить, что многофононный переход электрона (дырки) на примесный уровень может происходить с гораздо большей вероятностью, чем многофононное излучение при прямой рекомбинации (это связано с искажением решетки вблизи примесного центра при таком переходе).

Для ступенчатой рекомбинации можно написать уравнение для скорости рекомбинации, аналогичное (9.3). При наличии, например, доноров число свободных электронов, захватываемых донорами в единицу времени, пропорционально плотности электронов и плотности не занятых (положительно заряженных) доноров; число электронов, рекомбинирующих на донорах с дырками, пропорционально плотности занятых (нейтральных) доноров и плотности дырок.

Ступенчатая рекомбинация может идти при наличии нескольких уровней захвата на примесных центрах и с участием третьей частицы (оже-процессы).

Все эти случаи, на которых мы не можем останавливаться подробно, широко изучены как экспериментально, так и теоретически [2]).

---

[1]) Можно было бы в правой части (9.1) дополнительно учесть член, связанный с генерацией электронно-дырочных пар, например в результате поглощения света.

[2]) См., например, Пикус Г. Е., гл. 6; Paul R. Hableiterphysik.— Berlin.: VEB Verlag Tecknik, 1974, гл. 5, 6.

2. В том случае, когда концентрации электронов и дырок не являются равновесными, можно в большинстве случаев считать, что электроны и дырки в отдельности находятся в состоянии статистического равновесия. Это связано с тем, что характерное время между столкновениями для носителей тока в каждой из зон порядка 10^{-12} *сек*, в то время как среднее время для процессов рекомбинации меняется в широких пределах от 10^{-3} до 10^{-11} *сек*. Таким образом, электроны (дырки) приходят внутри зоны к равновесию гораздо быстрее, чем рекомбинируют.

Если полупроводник находится в состоянии статистического равновесия, то функции распределения электронов (2.1) и дырок (2.14) зависят от одного параметра — химического потенциала ζ.

Если электроны в зоне проводимости и дырки в валентной зоне в отдельности находятся в статистическом равновесии, то можно ввести различные химические потенциалы: для электронов — ζ_n и для дырок ζ_p, называемые *квазиуровнями Ферми*. Тогда неравновесные концентрации электронов и дырок в невырожденном случае, аналогично (2.15б) и (2.15в), равны

$$n = \frac{(2\pi m_n k_0 T)^{3/2}}{4\pi^3 \hbar^3} \exp \frac{\zeta_n}{k_0 T}, \tag{9.4}$$

$$p = \frac{(2\pi m_p k_0 T)^{3/2}}{4\pi^3 \hbar^3} \exp\left(-\frac{\varepsilon_G + \zeta_p}{k_0 T}\right). \tag{9.4а}$$

Очевидно, что при равновесии концентрации $n = n_0$, $p = p_0$ и квазиуровни Ферми совпадают с равновесным химическим потенциалом, т. е. $\zeta_n = \zeta_p = \zeta$. Наоборот, чем больше отличаются концентрации от равновесных, тем дальше отходят ζ_n и ζ_p от ζ. В самом деле, из (9.4), (9.4а), (2.15б) и (2.15в) следует, что

$$\zeta_n - \zeta = k_0 T \ln \frac{n}{n_0}, \quad \zeta - \zeta_p = k_0 T \ln \frac{p}{p_0}. \tag{9.5}$$

Рассмотрим теперь ситуацию, когда к полупроводнику с неравновесными концентрациями носителей тока приложено электрическое поле, так что текут электронный и дырочный токи.

Если приложенное к полупроводнику в данной точке $\boldsymbol{r}$ поле $\boldsymbol{E} = -\operatorname{grad}\Phi$ много меньше внутрикристаллического поля, т. е. много меньше нескольких сотен тысяч *в/см*, то энергетическая зонная структура полупроводника сохранится; при этом края зон приобретут в поле $\boldsymbol{E}$ одинаковый наклон, так что ширина запрещенной зоны остается постоянной, т. е. $\operatorname{grad}\varepsilon_G = 0$.

Энергия электрона проводимости в точке $\boldsymbol{r}$ в электрическом поле $\boldsymbol{E} = -\operatorname{grad}\Phi(\boldsymbol{r})$, где $\Phi(\boldsymbol{r})$ — электростатический потенциал, равна $\varepsilon_k - e\Phi(\boldsymbol{r})$, ε_k — «кинетическая энергия» электрона, зависящая от волнового вектора $\boldsymbol{k}$; аналогично энергия дырки равна $\varepsilon_{k'} + e\Phi(\boldsymbol{r})$, где кинетическая энергия дырки $\varepsilon_{k'} < 0$; для

стандартной зоны

$$\varepsilon_k = (\hbar^2/2m_n)\, k^2 \quad \text{и} \quad \varepsilon_{k'} = -(\hbar^2/2m_p)\, k'^2. \tag{9.6}$$

Если теперь в (2.1) и (2.14) заменить кинетические энергии электрона и дырки на полные энергии с учетом потенциальной энергии в поле $\Phi(\boldsymbol{r})$, то выражения (9.4) и (9.4а) приобретут вид

$$n(\boldsymbol{r}) = \frac{(2\pi m_n k_0 T)^{3/2}}{4\pi^3 \hbar^3} \exp \frac{\zeta_n(\boldsymbol{r}) + e\Phi(\boldsymbol{r})}{k_0 T}, \tag{9.7}$$

$$p(\boldsymbol{r}) = \frac{(2\pi m_p k_0 T)^{3/2}}{4\pi^3 \hbar^3} \exp \left(-\frac{\varepsilon_G + \zeta_p(\boldsymbol{r}) + e\Phi(\boldsymbol{r})}{k_0 T} \right). \tag{9.7а}$$

Очевидно, что в общем случае, если концентрации электронов $n(\boldsymbol{r})$ и дырок $p(\boldsymbol{r})$ зависят от координат, то квазиуровни Ферми тоже зависят от координат.

В трехмерном случае суммарный полевой и диффузионный токи для электронов и дырок, согласно (7.7), равны

$$\boldsymbol{j}_n = en\mu_n \boldsymbol{E} + k_0 T \mu_n \operatorname{grad} n, \tag{9.8}$$

$$\boldsymbol{j}_p = ep\mu_p \boldsymbol{E} - k_0 T \mu_p \operatorname{grad} p, \tag{9.8а}$$

где мы воспользовались соотношением Эйнштейна (7.11). Здесь μ_n и μ_p — подвижности электронов и дырок. Определим $\operatorname{grad} n(\boldsymbol{r})$ и $\operatorname{grad} p(\boldsymbol{r})$ из (9.7) и (9.7а) и учтем, что $\operatorname{grad}\Phi = -\boldsymbol{E}$, тогда из (9.8) и (9.8а) получим

$$\boldsymbol{j}_n = n\mu_n \operatorname{grad} \zeta_n, \tag{9.9}$$

$$\boldsymbol{j}_p = p\mu_p \operatorname{grad} \zeta_p. \tag{9.9а}$$

Таким образом, градиент квазиуровня Ферми учитывает как ток, связанный с градиентом электростатического потенциала (полевой ток), так и ток, связанный с градиентом концентрации (диффузионный ток).

ГЛАВА VII

ОПТИКА ПОЛУПРОВОДНИКОВ

§ 1. Дисперсионные соотношения Крамерса — Кронига

1. В этом параграфе мы рассмотрим на основе классической электродинамики некоторые вопросы оптики изотропных сред. Для оптически изотропной среды (например, кубического кристалла) в статическом или медленно меняющемся электромагнитном поле вектор электрической индукции

$$\boldsymbol{D} = \varepsilon \boldsymbol{E} = \boldsymbol{E} + 4\pi \boldsymbol{P}, \tag{1.1}$$

где $\boldsymbol{E}$ — напряженность среднего электрического поля в среде, $\boldsymbol{P}$ — вектор поляризации, ε — диэлектрическая проницаемость (диэлектрическая постоянная).

Соотношение (1.1) основано на предположениях об изотропности среды, линейной связи между $\boldsymbol{D}$ и $\boldsymbol{E}$, локальности и синхронности, которые предполагают, что $\boldsymbol{D}(\boldsymbol{r}, t)$ зависит от $\boldsymbol{E}(\boldsymbol{r}, t)$ в той же точке пространства, в тот же момент времени. Если связь между $\boldsymbol{D}(\boldsymbol{r}, t)$ и $\boldsymbol{E}(\boldsymbol{r}, t)$ не имеет локального и синхронного характера,— говорят о *пространственной* и *временной дисперсии*. Мы покажем, как меняется связь между $\boldsymbol{D}$ и $\boldsymbol{E}$ в быстро меняющемся электромагнитном поле, когда нарушается последнее условие — синхронность. Исследуем, может ли при этом сохраняться условие локальности взаимодействия, т. е. может ли в быстро меняющемся электромагнитном поле пространственная периодичность быть много большей размеров атомной структуры решетки. Под быстро меняющимся электромагнитным полем мы будем понимать такое, частота которого $\omega > \omega_p$ — частоты установления поляризации среды ($1/\omega_p \approx \tau_p$ — время установления поляризации). Наибольшая частота ω_p, связанная с перераспределением электронов в атоме, по порядку величины равна $\omega_p \approx v_a/a$, где v_a — скорость движения атомных электронов, a — размер атома. Наибольшая длина волны электромагнитного поля, соответствующая $\omega = \omega_p$, равна $\lambda = \frac{c}{\omega_p} \approx \frac{c}{v_a} \cdot a \gg a$, так как скорость v_a по крайней мере на два порядка меньше скорости света c.

Таким образом, и для быстрых электромагнитных полей частоты $\omega \geqslant \omega_P$ длина волны λ может оставаться большей атомных размеров, что позволяет пользоваться макроскопическими уравнениями электродинамики, т. е. локальным описанием явлений[1]).

Излагаемые ниже соображения относятся как к диэлектрикам, так и полупроводникам и металлам.

Предполагая, что связь между $\boldsymbol{D}$ и $\boldsymbol{E}$ линейна, как и в (1.1), но частота электромагнитного поля настолько велика, что значение $\boldsymbol{D}(t)$ определяется не только значением $\boldsymbol{E}(t)$ в тот же момент времени, но и предшествующими этому моменту времени значениями $\boldsymbol{E}(t')$ $(-\infty < t' \leqslant t)$, получим

$$\boldsymbol{D}(t) = \boldsymbol{E}(t) + \int\limits_{-\infty}^{t} f(t-t')\,\boldsymbol{E}(t')\,dt'. \tag{1.2}$$

Здесь $f(t)$ — функция, определяемая только свойствами среды. Выделение в правой части (1.2) слагаемого $\boldsymbol{E}(t)$ представляется, как мы увидим ниже, удобным. То, что значение $\boldsymbol{D}(t)$ определяется значениями $\boldsymbol{E}(t')$ в предшествующие моменты времени, согласуется с *принципом причинности*.

Интеграл в правой части (1.2) может быть записан в виде

$$\int\limits_{-\infty}^{\infty} f(t-t')\,\boldsymbol{E}(t')\,dt', \tag{1.2а}$$

если сделать дополнительное предположение, что $f(t-t') = 0$ при $t' > t$.

Перейдем к фурье-представлению функций $\boldsymbol{D}(t)$, $\boldsymbol{E}(t)$ и $f(t)$:

$$\boldsymbol{D}(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \boldsymbol{D}(\omega)\, e^{-i\omega t}\, d\omega, \tag{1.3}$$

$$\boldsymbol{E}(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \boldsymbol{E}(\omega)\, e^{-i\omega t}\, d\omega, \tag{1.3а}$$

$$f(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} f(\omega)\, e^{-i\omega t}\, d\omega. \tag{1.3б}$$

Здесь предполагается, что для $t > 0$ функция $f(t) = 0$. Обратные преобразования имеют вид

$$\boldsymbol{D}(\omega) = \int\limits_{-\infty}^{\infty} \boldsymbol{D}(t)\, e^{i\omega t}\, dt \tag{1.4}$$

и аналогично для $\boldsymbol{E}(\omega)$ и $f(\omega)$.

---

[1]) Отметим, что вблизи края экситонного поглощения и в случае $\lambda \gg a$ необходимо учитывать пространственную дисперсию.

Подставляя (1.3), (1.3а), (1.3б) в (1.2), получим

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \boldsymbol{D}(\omega)\, e^{-i\omega t}\, d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \boldsymbol{E}(\omega)\, e^{-i\omega t}\, d\omega + \\ + \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega'\, f(\omega)\, \boldsymbol{E}(\omega')\, e^{-i\omega(t-t')} e^{-i\omega' t'}.$$

Так как

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dt' e^{i(\omega-\omega')t'} = \delta(\omega-\omega'),$$

то

$$\int_{-\infty}^{\infty} \{\boldsymbol{D}(\omega) - [1+f(\omega)]\,\boldsymbol{E}(\omega)\}\, e^{-i\omega t}\, d\omega = 0.$$

Так как интеграл должен равняться нулю для всех значений t, то

$$\boldsymbol{D}(\omega) = \tilde{\varepsilon}(\omega)\, \boldsymbol{E}(\omega), \tag{1.5}$$

где

$$\tilde{\varepsilon}(\omega) = 1 + f(\omega) = 1 + \int_0^{\infty} f(t)\, e^{i\omega t}\, dt. \tag{1.5а}$$

Из (1.5) мы видим, что в случае быстро меняющегося электромагнитного поля связь между фурье-компонентами электрической индукции $\boldsymbol{D}$, напряженности поля $\boldsymbol{E}$ и диэлектрической проницаемостью $\tilde{\varepsilon}$ такая же, как для этих величин в стационарном случае (1.1).

Под *частотной дисперсией* диэлектрической проницаемости $\tilde{\varepsilon}(\omega)$ понимают ее зависимость от частоты ω.

Так как $f(t)$ — вещественная функция, то из (1.5а) следует, что диэлектрическая проницаемость $\tilde{\varepsilon}(\omega)$ комплексна, т. е.

$$\tilde{\varepsilon}(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega), \tag{1.6}$$

где $\varepsilon_1(\omega)$ — ее вещественная, а $\varepsilon_2(\omega)$ — мнимая часть.

Из (1.5а) следует, что $\tilde{\varepsilon}(\omega)$ при изменении знака ω переходит в $\tilde{\varepsilon}^*(\omega)$, т. е. $\varepsilon_1(-\omega) + i\varepsilon_2(-\omega) = \varepsilon_1(\omega) - i\varepsilon_2(\omega)$, откуда

$$\varepsilon_1(-\omega) = \varepsilon_1(\omega), \quad \varepsilon_2(-\omega) = -\varepsilon_2(\omega). \tag{1.6а}$$

Таким образом, вещественная часть диэлектрической проницаемости четная, а мнимая — нечетная функции ω. Как мы увидим ниже $\varepsilon_1(\omega)$ и $\varepsilon_2(\omega)$ просто связаны с оптическими константами среды: показателем преломления и коэффициентом поглощения.

Между $\varepsilon_1(\omega)$ и $\varepsilon_2(\omega)$ существуют важные соотношения (см. Приложение 19)

$$\varepsilon_1(\omega)-1=\frac{1}{\pi}\int\limits_{-\infty}^{\infty}\frac{\varepsilon_2(x)}{x-\omega}\,dx=\frac{2}{\pi}\int\limits_{0}^{\infty}\frac{x\varepsilon_2(x)}{x^2-\omega^2}\,dx, \tag{1.7}$$

$$\varepsilon_2(\omega)=-\frac{1}{\pi}\int\limits_{-\infty}^{\infty}\frac{\varepsilon_1(x)-1}{x-\omega}\,dx=-\frac{2\omega}{\pi}\int\limits_{0}^{\infty}\frac{\varepsilon_1(x)}{x^2-\omega^2}\,dx, \tag{1.7a}$$

называемые *дисперсионными соотношениями* или *соотношениями Крамерса—Кронига* (1927). Интегралы в (1.7) и (1.7а) надо понимать в смысле их главного значения [1]).

Вывод дисперсионных соотношений (1.7) и (1.7а) существенно основан на том, что функция $f(t-t')=0$ для $t'>t$, т. е. на принципе причинности. Соотношения (1.7) и (1.7а) могут быть использованы для обработки экспериментальных данных: так как $\varepsilon_1(\omega)$ и $\varepsilon_2(\omega)$, как мы уже указывали, непосредственно связаны с оптическими постоянными среды — показателем преломления и коэффициентом поглощения (см. ниже).

Дисперсионные соотношения (1.7) показывают, что если нам известна (например, из опыта) зависимость $\varepsilon_2(\omega)$ во всем интервале частот, то мы можем определить зависимость $\varepsilon_1(\omega)$ (и обратно). Конечно, практически нам никогда не известна зависимость $\varepsilon_2(\omega)$ во всем интервале частот. Однако если нас интересует $\varepsilon_1(\omega)$ для частоты ω, то вклад $\varepsilon_2(x)$ для x, далеких от ω, несуществен. Форма кривой $\varepsilon_1(\omega)$ в точке ω определяется значениями $\varepsilon_2(x)$ для x, близких ω. Это приводит к определенным соотношениям между кривыми $\varepsilon_1(\omega)$ и $\varepsilon_2(\omega)$ вблизи ω. Велицкий [2]) показал, что

пик показателя преломления $\leftrightarrow$ край поглощения, (1.8)

спад показателя преломления $\leftrightarrow$ пик поглощения. (1.9)

Эти соответствия обнаруживаются на опыте.

2. Еще Максвелл показал, что диэлектрическая постоянная среды равна квадрату показателя преломления. В предыдущем пункте мы показали, что диэлектрическая проницаемость при учете дисперсии становится комплексной величиной, и отметили, что ее вещественная и мнимая части связаны с оптическими характеристиками среды.

Введем комплексный показатель преломления

$$\tilde{n}=n+ik. \tag{1.10}$$

Для того чтобы определить физический смысл его вещественной части n и мнимой части k, положим, что он связан с комплекс-

[1]) Смирнов В. И., т. III, ч. 2, п. 26.

[2]) Velický B.—Czech. J. Phys., 1961, v. B11, p. 787.

ной диэлектрической проницаемостью $\tilde{\varepsilon}(\omega)$ соотношением Максвелла

$$\tilde{\varepsilon}(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega) = \tilde{n}^2 = n^2 - k^2 + i \cdot 2nk. \tag{1.11}$$

Отсюда следует, что

$$\varepsilon_1(\omega) = n^2 - k^2, \tag{1.12}$$

$$\varepsilon_2(\omega) = 2nk. \tag{1.12а}$$

Рассмотрим распространение плоской электромагнитной волны вдоль оси x в среде с комплексной диэлектрической проницаемостью $\tilde{\varepsilon}$; как мы сейчас увидим, в этом случае волновой вектор $\varkappa$ тоже должен быть комплексным. Для электрического поля волны имеем

$$\boldsymbol{E} = \boldsymbol{E}_0 \exp i(\varkappa x - \omega t). \tag{1.13}$$

Подставляя это выражение в волновое уравнение [1])

$$\nabla^2 \boldsymbol{E} = \frac{\tilde{\varepsilon}}{c^2} \ddot{\boldsymbol{E}}$$

с комплексной диэлектрической постоянной $\tilde{\varepsilon}$, получим $\tilde{\varkappa}^2 = (\omega/c)^2 \tilde{\varepsilon} = (\omega/c)^2 \tilde{n}^2$, или

$$\tilde{\varkappa} = \frac{\omega}{c} \tilde{n} = \frac{\omega}{c}(n + ik). \tag{1.13а}$$

Подставляя это значение $\tilde{\varkappa}$ в (1.13), получим

$$E = E_0 \exp\left(-\frac{\omega k}{c} x\right) \exp\left[iw\left(\frac{x}{c_1} - t\right)\right], \tag{1.13б}$$

где $c_1 = c/n$ — скорость света в среде. Мы видим, что вещественная часть комплексного показателя преломления n — обычный показатель преломления, а мнимая часть k, называемая *коэффициентом экстинкции*, определяет поглощение в среде (при $x = c/\omega k$, амплитуда электрического поля падает в $e \approx 2{,}718$ раза) [2]).

При распространении света вдоль x в поглощающей среде уменьшение интенсивности света $dI(x) = -\alpha I\, dx$, где α — коэффициент поглощения, откуда

$$I(x) = I_0 e^{-\alpha x}. \tag{1.14}$$

Здесь $I_0 = I(0)$. Так как $I \backsim E^2$, то из (1.14) и (1.13б) следует, что

$$\alpha = 2\omega k/c. \tag{1.15}$$

---

[1]) Тамм И. Е., § 100, (100.2).

[2]) Мы надеемся, что читателя не затруднит то, что мы здесь для показателя преломления и коэффициента экстинкции используем общепринятые обозначения n и k, хотя в других разделах книги они имеют смысл концентрации свободных носителей и волнового вектора электрона.

Из (1.12) и (1.12а) следует, что если поглощение мало, то $\varepsilon_1(\omega)\approx n^2$, и поглощение определяется мнимой частью диэлектрической проницаемости $\varepsilon_2(\omega)$.

Для экспериментального определения оптических констант n и k часто исследуют отражение света от поверхности тела. При нормальном падении света коэффициент отражения [1])

$$R=\left|\frac{\tilde{n}-1}{\tilde{n}+1}\right|=\frac{(n-1)^2+k^2}{(n+1)^2+k^2}. \tag{1.16}$$

Если поглощение мало́, то

$$R=\frac{(n-1)^2}{(n+1)^2}, \tag{1.16а}$$

и, следовательно, измерение коэффициента отражения позволяет определить показатель преломления n.

§ 2. Межзонное поглощение света, связанное с прямыми переходами

1. Поведение электронов в атомах и молекулах было в большой мере выяснено в результате оптических исследований, т. е. спектроскопии газов.

Интерпретация результатов оптических исследований в твердых телах, в частности полупроводниках, наталкивалась до недавнего времени на большие трудности, так как закономерности движения электронов в твердых телах более разнообразны и сложны, чем в атомах.

В большинстве полупроводников (германии, кремнии, антимониде индия, антимониде галлия, сульфидах, селенидах и теллуридах свинца и т. д.) оптическое поглощение света обладает следующим характером: до тех пор, пока частота света $\omega<\varepsilon_G/\hbar$, где ε_G — ширина запрещенной зоны, поглощение очень мало, но когда энергия фотона $\hbar\omega\geqslant\varepsilon_G$, поглощение быстро возрастает на несколько порядков; это связано с тем, что фотон передает свою энергию электронам валентной зоны, перебрасывая их в зону проводимости (*межзонное поглощение*).

Очевидно, что изучение спектра межзонного поглощения вблизи границы его быстрого роста (*края поглощения* [2])) дает информацию о структуре энергетического спектра электрона вблизи верхнего края валентной зоны и нижнего края зоны проводимости, что существенно для определения электрических свойств полупроводника.

---

[1]) Тамм И. Е., § 102.

[2]) Длина волны, соответствующая краю поглощения при $\varepsilon_G=1$ *эв*, равна $\lambda_G=2\pi c/\omega_G=2\pi c\hbar/\varepsilon_G=10^{-4}$ *см* $=1$ *мкм*.

Волновой вектор фотона, соответствующий краю поглощения, обычно много меньше волнового вектора электрона (дырки), образовавшегося в результате поглощения. В самом деле, волновой вектор фотона $\varkappa = 2\pi\lambda_G \approx 10^5\,см$, а волновой вектор электрона $k = \frac{1}{\hbar}\sqrt{2m(\hbar\omega - \varepsilon_G)}$, что при $\hbar\omega - \varepsilon_G \approx k_0 T$ и $T \sim 100°\,\mathrm{C}$ по порядку величины дает $k \approx 10^7\,см$.

Таким образом, при поглощении фотона электроном можно считать, что $\varkappa \ll k$. Так как при поглощении электроном фотона должен выполняться не только закон сохранения энергии, но и закон сохранения волнового вектора (квазиимпульса), то волновой вектор электрона $\boldsymbol{k}$ практически не меняется, т. е. электрон при поглощении фотона переходит из валентной зоны в зону проводимости в ту же точку зоны Бриллюэна (в ту же точку $\boldsymbol{k}$)— это так называемые *прямые переходы*.

Из рис. IV.28, *а* видно, что у германия максимум валентной зоны расположен в центре бриллюэновской зоны ($\boldsymbol{k} = 0$), а минимум зоны проводимости — на поверхности зоны Бриллюэна в точке L (см. рис. IV.22, $k_x = k_y = k_z = \pi/a$).

Опыт показывает, что в германии, наряду с прямыми переходами, наблюдаются *непрямые переходы* валентного электрона из центра бриллюэновской зоны в точку L зоны на ее поверхности.

Такое большое изменение волнового вектора электрона возможно только в том случае, если при поглощении фотона испускается или поглощается фонон с волновым вектором $q \approx \pi/a$.

Опыт показывает, что на частотную зависимость коэффициента поглощения света в ряде полупроводников влияет кулоновское взаимодействие электрона и дырки, образующихся при поглощении фотона.

При частотах, меньших частоты, соответствующей краю поглощения, можно наблюдать поглощение света электронами проводимости, связанное с переходами электрона внутри зоны проводимости. Следует иметь в виду, что свободный электрон, не взаимодействующий с искажениями правильной решетки, не может полностью поглотить фотон. В самом деле, законы сохранения энергии и импульса имеют в этом случае вид [1])

$$\frac{p^2}{2m^*} + \hbar\omega = \frac{p'^2}{2m^*}, \quad \boldsymbol{p} + \left(\frac{\overrightarrow{\hbar\omega}}{c}\right) = \boldsymbol{p}',$$

где $\boldsymbol{p} = m^*\boldsymbol{v}$ и $\boldsymbol{p}' = m^*\boldsymbol{v}'$ — импульсы электрона до и после поглощения фотона частоты ω, m^* — эффективная масса электрона, c — скорость света, $\left(\frac{\overrightarrow{\hbar\omega}}{c}\right)$ — импульс фотона. Подставляя $\boldsymbol{p}'$ из

[1]) Для простоты мы рассматриваем нерелятивистский случай и стандартную зону.

второго уравнения в первое, получим для косинуса угла между направлениями $\boldsymbol{p}$ и $\left(\frac{\hbar\vec{\omega}}{c}\right)$

$$\cos\vartheta = \frac{m^*c}{p}\left[1 - \frac{\hbar\omega}{2m^*c^2}\right] \approx \frac{m^*c}{p} = \frac{c}{v} \gg 1,$$

что не может выполняться ни для какого угла ϑ.

Поэтому для простых зон поглощение света свободными носителями возможно лишь при одновременном излучении или поглощении фонона или при передаче импульса примеси.

Важная информация о полупроводниках может быть получена при изучении поглощения света в них в электрическом и особенно магнитном поле.

2. Электромагнитное поле световой волны в вакууме может быть описано посредством *векторного потенциала* $\boldsymbol{A}(\boldsymbol{r}, t)$ и *скалярного потенциала* $\varphi(\boldsymbol{r}, t)$ [1]. Электромагнитные потенциалы $\boldsymbol{A}(\boldsymbol{r}, t)$ и $\varphi(\boldsymbol{r}, t)$ определены с точностью до градиентного преобразования, что позволяет положить скалярный потенциал $\varphi(\boldsymbol{r}, t) \equiv 0$; тогда электрическое поле $\boldsymbol{E}(\boldsymbol{r}, t)$ и магнитное поле $\boldsymbol{H}(\boldsymbol{r}, t)$ равны

$$\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}, \tag{2.1}$$

$$\boldsymbol{H} = \operatorname{rot} \boldsymbol{A}. \tag{2.2}$$

Мы можем векторный потенциал $\boldsymbol{A}(r, t)$ дополнительно подчинить условию Лоренца

$$\operatorname{div} \boldsymbol{A} = 0. \tag{2.3}$$

В этом случае $\boldsymbol{A}(r, t)$ удовлетворяет волновому уравнению.

Гамильтониан электрона с зарядом $-e$ в периодическом поле кристалла $V(\boldsymbol{r})$ и электромагнитном поле световой волны имеет вид (VI.5.12)

$$\hat{\mathcal{H}} = \frac{1}{2m}\left[\hat{\boldsymbol{p}} + \frac{e}{c}\boldsymbol{A}(\boldsymbol{r}, t)\right]^2 + V(\boldsymbol{r}), \tag{2.4}$$

где $\hat{\boldsymbol{p}} = -i\hbar\nabla$. Далее

$$\left[\hat{\boldsymbol{p}} + \frac{e}{c}\boldsymbol{A}\right]^2 = \hat{\boldsymbol{p}}^2 + \frac{e}{c}\hat{\boldsymbol{p}}\boldsymbol{A} + \frac{e}{c}\boldsymbol{A}\hat{\boldsymbol{p}} + \frac{e^2}{c^2}\boldsymbol{A}^2 = \hat{\boldsymbol{p}}^2 + 2\frac{e}{c}\boldsymbol{A}\hat{\boldsymbol{p}} + \frac{e^2}{c^2}\boldsymbol{A}^2, \tag{2.5}$$

где мы воспользовались перестановочными соотношениями [2]: $\hat{\boldsymbol{p}}\boldsymbol{A} = \boldsymbol{A}\hat{\boldsymbol{p}} - i\hbar \operatorname{div} \boldsymbol{A}$ и условием (2.3). Отношение третьего члена ко второму в правой части (2.5) по порядку величины равно

$$\frac{e}{c}\frac{|\boldsymbol{A}|}{p} = \frac{e|\boldsymbol{E}_0|}{\omega p} \approx \frac{e}{\omega p}\left(\frac{8\pi S}{c}\right)^{1/2}, \tag{2.6}$$

где мы использовали (2.10); $S = \frac{c}{8\pi}|\boldsymbol{E}_0|^2$ — плотность светового

[1] Тамм И. Е., §§ 94, 96.
[2] Блохинцев Д. И., § 24.

потока (вектор Пойнтинга). Для мощного светового потока $S=1$ *вт/см*^{2} (десятикратный поток солнечного излучения на границе атмосферы), частоты света $\omega=10^{14}$ *сек*$^{-1}$, нормальной температуры $T\approx 300°$K (для которой средний импульс $p\approx\hbar k\approx 10^{-20}$ *г·см·сек*$^{-1}$), отношение (2.6) порядка 10^{-4}, поэтому в (2.5) можно пренебречь последним слагаемым.

Таким образом, в гамильтониане (2.4) возмущение, связанное с действием световой волны, равно

$$\mathscr{H}'=\frac{e}{mc}\boldsymbol{A}\hat{\boldsymbol{p}}=-\frac{ie\hbar}{mc}\boldsymbol{A}\nabla. \tag{2.7}$$

Для плоской монохроматической волны вектор-потенциал

$$\boldsymbol{A}(\boldsymbol{r},t)=A_0\boldsymbol{e}\exp i(\boldsymbol{\varkappa r}-\omega t), \tag{2.8}$$

где A_0—амплитуда (вообще говоря, комплексная), $\boldsymbol{e}$—единичный вектор, ω и $\boldsymbol{\varkappa}$—частота и волновой вектор электромагнитной волны.

Электрическое поле, как следует из (2.1), равно

$$\boldsymbol{E}=-\frac{1}{c}\frac{\partial\boldsymbol{A}}{\partial t}=\frac{i\omega}{c}\boldsymbol{A}, \tag{2.9}$$

так что вектор поляризации $\boldsymbol{e}$ направлен вдоль электрического поля. Амплитуда электрического поля

$$|E_0|=\frac{\omega}{c}|A_0|. \tag{2.10}$$

Из (2.8) и (2.3) следует, что

$$\operatorname{div}\boldsymbol{A}=iA_0\exp i(\boldsymbol{\varkappa r}-\omega t)\cdot(\boldsymbol{e\varkappa})=0, \tag{2.11}$$

т. е. $\boldsymbol{\varkappa}$ перпендикулярно $\boldsymbol{e}$ или $\boldsymbol{E}$ (электромагнитные волны поперечны).

Для того чтобы выразить модуль амплитуды вектор-потенциала $|A_0|$ через энергию фотонов в 1 *см*^{3} $N\hbar\omega$ ($\hbar\omega$—энергия фотона, N—их число в 1 *см*^{3}), перейдем к вещественным значениям полей[1]), положив вектор-потенциал

$$A(\boldsymbol{r},t)=A_0\exp i(\boldsymbol{\varkappa r}-\omega t)+A_0^*\exp[-i(\boldsymbol{\varkappa r}-\omega t)]=2|A_0|\cos(\boldsymbol{\varkappa r}-\omega t), \tag{2.12}$$

где мы, не ограничивая общности, считаем, что $A_0=A_0^*=|A_0|$. Вещественное значение напряженности электрического поля равно

$$E=-\frac{1}{c}\frac{\partial A}{\partial t}=2\frac{\omega}{c}|A_0|\sin(\boldsymbol{\varkappa r}-\omega t).$$

Для немагнитных оптически изотропных сред, к которым относятся кубические кристаллы (Ge, Si, InSb и др.), средняя по

[1]) Известно, что непосредственное использование комплексных значений полей в выражениях квадратичных относительно полей (например, плотности энергии) недопустимо (см., например, И. Е. Тамм, стр. 473).

времени плотность энергии в электромагнитной волне равна (n — показатель преломления) [1])

$$\frac{1}{8\pi}[n^2\overline{E^2}+\overline{H^2}]=\frac{1}{4\pi}n^2\overline{E^2}=\frac{1}{4\pi}\cdot 4\left(\frac{\omega}{c}\right)^2|A_0|^2\overline{\sin^2(\varkappa r-\omega t)}=$$

$$=\frac{1}{2\pi}\left(\frac{\omega}{c}\right)^2|A_0|^2.$$

При этом мы воспользовались тем, что в плоской гармонической волне средняя электрическая энергия $n^2\overline{E^2}$ равна средней магнитной энергии $\overline{H^2}$ и что $\overline{\sin^2(\varkappa r-\omega t)}={}^1/_2$.

Приравнивая полученное выше выражение плотности энергии фотонов $N\hbar\omega$, получим

$$|A_0|=\sqrt{2\pi N\hbar\omega}/\varkappa. \tag{2.12a}$$

Здесь волновое число $\varkappa=\omega/v$, где фазовая скорость $v=c/n$. Для энергии возмущения (2.7) имеем

$$\hat{\mathcal{H}}'=-\frac{i\hbar e}{mc}A_0\exp[-i(\omega t-\varkappa r)]\cdot(e\nabla), \tag{2.13}$$

где $|A_0|$ равно (2.12a) [2]).

Состояние электрона в зоне проводимости и валентной зоне описывается блоховской волновой функцией

$$\psi_{nk}(r)=\exp\left[-\frac{i}{\hbar}\varepsilon_n(k)t\right]\cdot u_{nk}(r)\exp ikr, \tag{2.14}$$

где n — номер зоны, k — приведенный волновой вектор электрона, $\varepsilon_n(k)$ — энергия электрона, $u_{nk}(r)$ — периодическая функция с периодами решетки.

Матричный элемент энергии возмущения (2.13), связанный с поглощением фотона, для перехода $(n_1, k_1)\longrightarrow(n_2, k_2)$ равен

$$\langle n_2, k_2|\hat{\mathcal{H}}'|n_1, k_1\rangle=\int d^3r\psi^*_{n_2k_2}\hat{\mathcal{H}}'\psi_{n_1k_1}=$$

$$=-\frac{i\hbar e}{mc}A_0\int d^3r\,[u^*_{n_2k_2}\exp[-i(k_2-\varkappa)r]\cdot e\nabla u_{n_1k_1}\exp ik_1r]\times$$

$$\times\exp\frac{i}{\hbar}(\varepsilon_{n_2}(k_2)-\varepsilon_{n_1}(k_1)-\hbar\omega)t=\mathcal{P}_{12}\exp\frac{i}{\hbar}(\varepsilon_{n_2}(k_2)-\varepsilon_{n_1}(k_1)-\hbar\omega)t, \tag{2.15}$$

где $\mathcal{P}_{12}$ — не зависящая от времени часть матричного элемента.

---

[1]) Тамм И. Е., § 92.

[2]) Строго говоря, в (2.13) следовало бы для $A(r, t)$ использовать двухчленное вещественное выражение (2.12), второе слагаемое которого, пропорциональное $\exp i\omega t$, учитывает испускание фотонов. Ввиду малого количества возбужденных при межзонном поглощении света электронов этим процессом можно пренебречь; это условие не всегда выполняется (лазеры).

Подынтегральное выражение равно

$$[\ldots]\exp i(\ldots)\equiv[u^*_{n_2k_2}e\nabla u_{n_1k_1}+i(ek_1)u^*_{n_2k_2}u_{n_1k_1}]\exp i(k_1+\varkappa-k_2)r. \tag{2.16}$$

Положим радиус-вектор $r=a_m+r'$, где a_m—вектор узла решетки m. В силу периодичности функций $u_{nk}(r)$ интеграл (2.15) примет вид

$$\sum_m \exp i(k_1+\varkappa-k_2)a_m\int[\ldots]\exp i(k_1+\varkappa-k_2)r'\,d^3r'. \tag{2.17}$$

Здесь $[\ldots]$ обозначено выражение в квадратной скобке в (2.16), а интегрирование по d^3r' ведется по объему элементарной ячейки Ω_0. Из (П. 6.4) следует, что

$$\sum_m \exp(ik_1+\varkappa-k_2)a_m = N\delta_{0,k_1+\varkappa-k_2}=$$

$$=\frac{V}{\Omega_0}\delta_{0,k_1+\varkappa-k_2}=\frac{1}{\Omega_0}\delta_{0,k_1+\varkappa-k_2}. \tag{2.18}$$

Здесь $\delta_{0,k_1+\varkappa-k_2}$—символ Кронекера, N—число элементарных ячеек объема Ω_0 в основной области кристалла $V=1\,см^3$; мы положили в (П. 6.4) вектор обратной решетки $b_g=0$, так как k_1 и k_2 лежат в первой зоне Бриллюэна, а $\varkappa\ll k_1$ и $\varkappa\ll k_2$. (2.18) выражает закон сохранения волнового вектора (квазиимпульса) и в силу малости $\varkappa$ сводится к равенствам

$$\varkappa\approx 0,\quad k_1=k_2=k. \tag{2.19}$$

Условие (2.19) позволяет положить в (2.17) экспоненту равной единице[1]); тогда интеграл от второго слагаемого в $[\ldots]$ равен нулю в силу ортогональности функций $u_{n_1k_1}$ и $u_{n_2k_2}$.

Таким образом, не зависящий от времени матричный элемент $\mathcal{P}_{12}$ в (2.15) равен

$$\mathcal{P}_{vc}=-\frac{i\hbar e}{mc}A_0\frac{1}{\Omega_0}\int\limits_{\Omega_0}d^3r\,u^*_{ck}(r)\,e\nabla u_{vk}(r)=A_0\frac{e}{mc}(ep_{cv}), \tag{2.20}$$

где мы заменили индексы n_1 и n_2 на v (валентная зона) и c (зона проводимости), положили $k_1=k_2=k$ и опустили штрих у переменной интегрирования r'; здесь матричный элемент перехода для импульса

$$p_{cv}=-i\hbar\frac{1}{\Omega_0}\int\limits_{\Omega_0}d^3r\,u^*_{ck}(r)\,\nabla u_{vk}(r) \tag{2.20а}$$

(множитель $1/\Omega_0$ введен для удобства).

Число переходов $v\to c$ в единицу времени в единице объема, для которых выполняется закон сохранения энергии ($\varepsilon_c=\varepsilon_v+\hbar\omega$)

[1]) Это соответствует дипольному приближению при оптических переходах в атоме.

и закон сохранения квазиимпульса или волнового вектора ($\boldsymbol{k}_1 = \boldsymbol{k}_2$), равно (см. (П. 20.10а))

$$\mathscr{W}_{vc} = \frac{2\pi}{\hbar} \int |\mathscr{P}_{vc}|^2 \delta[(\varepsilon_c(\boldsymbol{k}) - \varepsilon_v(\boldsymbol{k}) - \hbar\omega] \frac{2}{(2\pi)^3} d^3k. \quad (2.21)$$

Здесь $[2/(2\pi)^3]\, d^3k$ — число электронных состояний (с учетом спина) в области d^3k; интегрирование ведется по d^3k с учетом закона сохранения энергии. Мы предполагаем, что зона 1 (валентная) плотностью заполнена электронами, а зона 2 (проводимости) целиком свободна; такое предположение хорошо выполняется, если $k_0 T \ll \varepsilon_G$. Очевидно, что число поглощенных в 1 *см*^{3} в 1 *сек* фотонов равно числу переходов $\mathscr{W}_{vc}$ (2.21). Если в полупроводнике распространяется плоская световая волна, то плотность потока фотонов (число фотонов, проходящих сквозь 1 *см*^{2} в 1 *сек*) равна $Nv = Nc/n$, где N — число фотонов в 1 *см*^{3}, а фазовая скорость $v = c/n$ — скорость света в вакууме, деленная на показатель преломления.

Коэффициент поглощения, обусловленный прямыми межзонными переходами, равен [1])

$$\alpha = \frac{\mathscr{W}_{vc}}{Nv} = \frac{1}{\pi} \frac{e^2}{m^2 cn\omega} \int d^3k\, |\boldsymbol{e}\boldsymbol{p}_{cv}|^2 \delta[\varepsilon_c(\boldsymbol{k}) - \varepsilon_v(\boldsymbol{k}) - \hbar\omega)]. \quad (2.22)$$

Мы воспользовались значением $|A_0|^2$ (2.12а) и положили волновой вектор фотона $\varkappa = \omega/v = \omega n/c$.

Пусть минимум в зоне проводимости и максимум в валентной зоне расположены в одной и той же точке зоны Бриллюэна $\boldsymbol{k}_0$. Такую зонную структуру имеют полупроводники PbTe, PbSe и PbS. Точки соответствующих экстремумов расположены вдоль осей ⟨111⟩ на границах зоны Бриллюэна (см. рис. IV.22, точки L).

Если поверхности постоянной энергии электрона и дырки вблизи своих экстремумов — эллипсоиды с параллельными осями, что также имеет место для указанных выше полупроводников, то

$$\varepsilon_c(\boldsymbol{k}) = \varepsilon_c(\boldsymbol{k}_0) + \frac{\hbar^2}{2}\left[\frac{(k_1 - k_{01})^2}{m_{c1}} + \frac{(k_2 - k_{02})^2}{m_{c2}} + \frac{(k_3 - k_{03})^2}{m_{c3}}\right], \quad (2.23)$$

$$\varepsilon_v(\boldsymbol{k}) = \varepsilon_v(\boldsymbol{k}_0) - \frac{\hbar^2}{2}\left[\frac{(k_1 - k_{01})^2}{m_{v1}} + \frac{(k_2 - k_{02})^2}{m_{v2}} + \frac{(k_3 - k_{03})^2}{m_{v3}}\right]. \quad (2.23а)$$

Здесь оси 1, 2, 3 направлены вдоль главных осей эллипсоидов, m_{ci}^{-1} и m_{vi}^{-1} — компоненты тензоров обратных эффективных масс электронов в зоне проводимости и дырок в валентной зоне. Аргумент δ-функции в (2.22) равен

$$\varepsilon_c(\boldsymbol{k}) - \varepsilon_v(\boldsymbol{k}) - \hbar\omega = \frac{\hbar^2}{2}\left[\frac{(k_1 - k_{01})^2}{\mu_1} + \frac{(k_2 - k_{02})^2}{\mu_2} + \frac{(k_3 - k_{03})^2}{\mu_3}\right] - (\hbar\omega - \varepsilon_0), \quad (2.24)$$

[1]) Очевидно, что это определение совпадает с (1.14).

где

$$\varepsilon_0 = \varepsilon_c(\boldsymbol{k}_0) - \varepsilon_v(\boldsymbol{k}_0), \tag{2.24а}$$

а приведенные компоненты тензора обратных эффективных масс

$$\frac{1}{\mu_i} = \frac{1}{m_{ci}} + \frac{1}{m_{vi}}. \tag{2.24б}$$

При значениях $\boldsymbol{k}$, близких к $\boldsymbol{k}_0$,

$$\boldsymbol{ep}_{cv}(\boldsymbol{k}) = \boldsymbol{ep}_{cv}(\boldsymbol{k}_0) + \left[\frac{\partial}{\partial \boldsymbol{k}} \boldsymbol{ep}_{cv}(\boldsymbol{k})\right]_{\boldsymbol{k}=\boldsymbol{k}_0} (\boldsymbol{k} - \boldsymbol{k}_0). \tag{2.25}$$

Если $\boldsymbol{ep}_{cv}(\boldsymbol{k}_0) \neq 0$, переход называется *разрешенным*. Ограничиваясь в этом случае в (2.25) первым членом, получим из (2.22), (2.24) и (2.25)

$$\alpha_{\text{раз}} = \frac{1}{\pi} \frac{e^2}{m^2 c n \omega} \sum_{\boldsymbol{k}_0} |\boldsymbol{ep}_{cv}(\boldsymbol{k}_0)|^2 \times$$

$$\times \int \delta \left\{ \frac{\hbar^2}{2} \left(\frac{(k_1 - k_{01})^2}{\mu_1} + \frac{(k_2 - k_{02})^2}{\mu_2} + \frac{(k_3 - k_{03})^2}{\mu_3} \right) - (\hbar\omega - \varepsilon_0) \right\} dk_1\, dk_2\, dk_3,$$

где $\sum_{\boldsymbol{k}_0}$ — сумма значений $|\boldsymbol{ep}_{cv}(\boldsymbol{k}_0)|^2$ для всех точек, соответствующих рассматриваемым экстремумам в зоне Бриллюэна.

Переходя к переменным интегрирования $k_i' = (k_i - k_{0i})/\sqrt{\mu_i}$, получим

$$\alpha_{\text{раз}} = \frac{1}{\pi} \frac{e^2}{m^2 c n \omega} \sum_{\boldsymbol{k}_0} |\boldsymbol{ep}_{cv}(k_0)|^2 \sqrt{\mu_1 \mu_2 \mu_3} \times$$

$$\times \int \delta \left\{ \frac{1}{2} \hbar^2 k'^2 - (\hbar\omega - \varepsilon_0) \right\} 4\pi k'^2\, dk',$$

где мы положили $dk_1'\, dk_2'\, dk_3' = 4\pi k'^2\, dk'$. Вводя переменную интегрирования $x = \hbar^2 k'^2/2$ и используя δ-функцию, получим

$$\alpha_{\text{раз}} = \frac{4\sqrt{2}\, e^2}{m^2 c n \omega} \sum_{\boldsymbol{k}_0} |\boldsymbol{ep}_{cv}(\boldsymbol{k}_0)|^2 \frac{\sqrt{\mu_1 \mu_2 \mu_3}}{\hbar^3} (\hbar\omega - \varepsilon_0)^{1/2}. \tag{2.26}$$

Очевидно, что для $\hbar\omega < \varepsilon_0$, т. е. ниже порога поглощения $\alpha_{\text{раз}}$ равен нулю.

В дальнейшем (§ 4) мы увидим, что $\alpha_{\text{раз}}$ определяется вещественной частью удельной электропроводности σ, которая в кубических кристаллах изотропна, поэтому $\sum_{\boldsymbol{k}_0} |\boldsymbol{ep}_{cv}(\boldsymbol{k}_0)|^2$ должна быть изотропна; в этом можно было бы убедиться прямым расчетом, учитывая взаимную ориентацию осей $\langle 111 \rangle$, вдоль которых направлены $\boldsymbol{k}_0$. Однако нетрудно видеть, что изотропный инвариант в этом случае должен равняться

$$\sum_{\boldsymbol{k}_0} |\boldsymbol{ep}_{cv}(\boldsymbol{k}_0)|^2 = 4 \frac{p_{cvx}^2 + p_{cvy}^2 + p_{cvz}^2}{3}. \tag{2.26а}$$

В случае стандартной зоны в точке $\boldsymbol{k}_0 = 0$ (2.26) приобретает вид

$$\alpha_{\text{раз}} = \frac{2e^2}{m^2 cn\omega} |\boldsymbol{ep}_{cv}(0)|^2 \left(\frac{2\mu}{\hbar^2}\right)^{3/2} (\hbar\omega - \varepsilon_0)^{1/2}, \qquad (2.26б)$$

где $\mu = \mu_1 = \mu_2 = \mu_3$ и $(1/\mu) = (1/m_c) + (1/m_v)$ и $\varepsilon_0 = \varepsilon_c(0) - \varepsilon_v(0)$. Для того чтобы не усложнять изложения, рассмотрим для запрещенных переходов случай стандартной зоны с экстремумами в точке $\boldsymbol{k}_0 = 0$. Основные зависимости коэффициента поглощения и, в частности, его зависимость от частоты света получаются при этом такими же, как и в случае нестандартной зоны.

Для запрещенных переходов $\boldsymbol{ep}_{cv}(0) = 0$, поэтому разложение (2.25) начинается со второго члена, квадрат модуля которого равен

$$\left|\frac{\partial}{\partial \boldsymbol{k}}[\boldsymbol{ep}_{cv}(0)]\boldsymbol{k}\right|^2 = \left|\frac{\partial}{\partial \boldsymbol{k}}[\boldsymbol{ep}_{cv}(0)]\right|^2 k^2 \cos^2\vartheta.$$

Здесь ϑ — угол между векторами $\frac{\partial}{\partial \boldsymbol{k}}[\boldsymbol{ep}_{cv}(0)]$ и $\boldsymbol{k}$. Аргумент δ-функции в (2.22) в случае стандартной зоны равен $(\hbar^2 k^2/2\mu) - (\hbar\omega - \varepsilon_0)$.

Подставляя приведенные выше выражения в (2.22) и полагая в полярных координатах $d^3k = \sin\vartheta\, d\vartheta\, d\varphi k^2\, dk$ (полярная ось совпадает с вектором $\frac{\partial}{\partial \boldsymbol{k}}[\boldsymbol{ep}_{cv}(0)]$, получим

$$\alpha_{\text{зап}} = \frac{1}{\pi}\frac{e^2}{m^2 cn\omega}\left|\frac{\partial}{\partial \boldsymbol{k}}[\boldsymbol{ep}_{cv}(0)]\right|^2 \times$$
$$\times \int \delta\left[\frac{\hbar^2 k^2}{2\mu} - (\hbar\omega - \varepsilon_0)\right] k^2 \cos^2\vartheta \sin\vartheta\, d\vartheta\, d\varphi\, k^2\, dk.$$

Интегрирование по углам дает множитель $4\pi/3$, а интегрирование по k производится совершенно аналогично предыдущему случаю. В результате для коэффициента получим

$$\alpha_{\text{зап}} = \frac{2e^2}{3m^2 cn\omega}\left(\frac{2\mu}{\hbar^2}\right)^{5/2}\left|\frac{\partial}{\partial \boldsymbol{k}}[\boldsymbol{ep}_{cv}(0)]\right|^2 (\hbar\omega - \varepsilon_0)^{3/2}. \qquad (2.27)$$

Сравнивая это выражение с (2.26), мы видим, что в случае разрешенных переходов $\alpha_{\text{раз}} \sim (\hbar\omega - \varepsilon_0)^{1/2}$, в то время как в случае запрещенных переходов $\alpha_{\text{зап}} \sim (\hbar\omega - \varepsilon_0)^{3/2}$; это связано с тем, что в последнем случае под знаком интеграла (2.22) появляется лишний множитель k^2.

Из (1.12а) и (1.15) следует, что мнимая часть диэлектрической проницаемости $\varepsilon_2(\omega) = (nc/\omega)\alpha$. Подставляя сюда вместо α значения (2.26) или (2.27), мы получим величину $\varepsilon_2(\omega)$ для разрешенных и запрещенных межзонных переходов, не содержащую показателя преломления n.

Опыт показывает, однако, что зависимость коэффициента поглощения света от частоты ω для разрешенных и запрещенных

межзонных переходов носит, вообще говоря, более сложный характер, чем это следует из (2.26) или (2.27). Как было выяснено, на коэффициенте поглощения света может существенно сказываться кулоновское взаимодействие электрона и дырки, образующихся при поглощении фотона, т. е. свободные и связанные экситонные состояния. Учет влияния кулоновского взаимодействия электрона и дырки на поглощение света при прямых и непрямых межзонных переходах был рассмотрен Р. Эллиотом [1]).

3. В гл. V, § 3, п. 1 мы рассмотрели элементарную теорию экситонов большого радиуса.

Пусть экситон образовался в результате поглощения электроном в валентной зоне фотона. Закон сохранения волнового вектора при образовании экситона, ввиду малости волнового вектора фотона $\varkappa$, сводится к тому, что волновые векторы электрона $\boldsymbol{k}_n$ и дырки $\boldsymbol{k}_p$ равны по величине и противоположны по направлению, так что волновой вектор экситона $\boldsymbol{K}=\boldsymbol{k}_n+\boldsymbol{k}_p\approx 0$, поэтому кинетическая энергия образовавшегося экситона, как целого, (V.3.6) $W=0$.

Мы видим, что полная энергия экситона (V.3.5б)

$$\mathscr{E}=\varepsilon_0+\varepsilon, \tag{2.28}$$

где мы обозначили разность $\varepsilon_c(0)-\varepsilon_v(0)=\varepsilon_0$, которая в случае прямых переходов может не совпадать с ε_G.

Уравнение Шредингера, описывающего относительное движение электрона и дырки, имеет вид (V.3.5а)

$$-\frac{\hbar^2}{2\mu}\nabla_r^2\varphi(\boldsymbol{r})-\frac{e^2}{n^2 r}\varphi(\boldsymbol{r})=\varepsilon\varphi(\boldsymbol{r}). \tag{2.29}$$

Здесь μ — приведенная масса электрона и дырки, а статическая диэлектрическая постоянная ε_0 положена равной квадрату показателя преломления ($\varepsilon_0=n^2$); это уравнение совпадает с уравнением для водородоподобного атома с зарядом ядра e/n и массой электрона μ. Для связанных состояний экситона

$$\mathscr{E}=\varepsilon_0+\varepsilon=\varepsilon_0-\frac{\varepsilon_{\text{экс}}}{\nu^2}, \tag{2.30}$$

где главное квантовое число $\nu=1, 2, 3, \ldots$, а

$$\varepsilon_{\text{экс}}=\mu e^4/2\hbar^2 n^4. \tag{2.30а}$$

Наряду со связанными состояниями экситона можно рассмотреть несвязанные состояния, которым соответствуют в (2.29) положительные значения энергии ε. Решения уравнения (2.29) для $\varepsilon>0$ известны из теории водородоподобного атома. В этом случае относительное состояние движения образовавшихся электрона и дырки

[1]) Elliot R. J.—Phys. Rev., 1957, v. 108, p. 1384.

характеризуется волновым вектором $|\boldsymbol{k}|=\frac{1}{\hbar}\sqrt{2\mu\varepsilon}$, где μ—приведенная масса. При выводе (2.26) и (2.27) мы предполагали, что электрон и дырка после их образования двигаются независимо и каждый из них описывается плоской волной. На самом деле при учете корреляции их движения, коэффициент поглощения α увеличивается настолько, насколько увеличивается плотность вероятности состояния электрона на месте дырки (или наоборот — дырки на месте электрона) по сравнению с состоянием, описываемым плоской волной.

Можно показать[1]), что для несвязанных состояний отношение квадрата модуля волновой функции $\varphi(\boldsymbol{r})$ в нуле к квадрату модуля плоской волны, описывающей свободное движение, равно

$$\frac{|\varphi(0)|^2}{|\psi_k|^2}=\frac{\pi\beta\exp\pi\beta}{\operatorname{sh}\pi\beta}, \tag{2.31}$$

где

$$\beta=\left(\frac{2\mu}{\hbar^2}\right)^{1/2}\frac{\varepsilon_{\text{экс}}^{1/2}}{k}=\frac{1}{ka_B} \tag{2.31а}$$

и боровский радиус

$$a_B=\hbar^2n^2/\mu e^2. \tag{2.31б}$$

Представляется естественным (и это подтверждается детальным анализом Эллиотта), что для учета кулоновского взаимодействия электрона и дырки необходимо в подынтегральные выражения для $\alpha_{\text{раз}}$ и $\alpha_{\text{зап}}$ ввести в качестве весового множителя $|\varphi(0)|^2/|\psi_k|^2$ (2.31).

Таким образом, для стандартной зоны получим

$$\alpha_{\text{раз}}^{\text{кул}}=\frac{1}{\pi}\frac{e^2}{m^2c\omega}|\boldsymbol{ep}_{cv}(0)|^2\int\delta\left[\frac{\hbar^2k^2}{2\mu}-(\hbar\omega-\varepsilon_0)\right]4\pi k^2\frac{\frac{\pi}{a_Bk}\exp\frac{\pi}{a_Bk}}{\operatorname{sh}\frac{\pi}{a_Bk}}\,dk.$$

Используя, как и в предыдущем случае, для интегрирования δ-функцию, получим аналогично (2.26а)

$$\alpha_{\text{раз}}^{\text{кул}}=\frac{2\pi e^2}{nm^2\omega c}\left(\frac{2\mu}{\hbar^2}\right)^{3/2}|\boldsymbol{ep}_{cv}(0)|^2\varepsilon_{\text{экс}}^{1/2}\frac{e^z}{\operatorname{sh}z}, \tag{2.32}$$

где

$$z=\pi\sqrt{\frac{\varepsilon_{\text{экс}}}{\hbar\omega-\varepsilon_0}}. \tag{2.32а}$$

Вблизи края поглощения $\hbar\omega-\varepsilon_0\ll\varepsilon_{\text{экс}}$, т. е. $z\gg1$; влияние кулоновского взаимодействия велико. В пределе, когда $\hbar\omega\to\varepsilon_0$, т. е. $z\to\infty$, коэффициент поглощения не стремится к нулю,

[1]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика, с. 646. Следует учесть, что формула (136.11) записана в атомных единицах ($a_B=1$).

как $\alpha_{\text{раз}}$ (2.26), а, как следует из (2.32), стремится к величине

$$\alpha_{\text{раз}}^{\text{кул}}\Big|_{\hbar\omega\to\varepsilon_0}=\frac{4\pi e^2}{nm^2c\omega}\left(\frac{2\mu}{\hbar^2}\right)^{3/2}|\boldsymbol{ep}_{cv}(0)|^2\varepsilon_{\text{экс}}^{1/2}. \qquad (2.33)$$

Далеко от края поглощения, когда $\hbar\omega-\varepsilon_0\gg\varepsilon_{\text{экс}}$, кулоновское взаимодействие менее существенно. Однако даже при относительно больших значениях кинетической энергии электрона и дырки, например при $\hbar\omega-\varepsilon_0=\pi^2\varepsilon_{\text{экс}}\approx 10\varepsilon_{\text{экс}}$, т. е. $z=1$, кулоновское взаимодействие между ними в 4,6 раза увеличивает коэффициент поглощения $\alpha_{\text{раз}}$

$$\alpha_{\text{раз}}^{\text{кул}}/\alpha_{\text{раз}}=2ze^z/\operatorname{sh} z=2e/\operatorname{sh} 1=4{,}6.$$

Эллиотт показал, что для поглощения, связанного с запрещенными переходами в состояние $\varphi(\boldsymbol{r})$, весовой множитель под знаком интеграла для $\alpha_{\text{зап}}^{\text{кул}}$ равен $|\partial\varphi/\partial\boldsymbol{r}|^2_{\boldsymbol{r}=0}$. Расчет, аналогичный предыдущему, дает

$$\alpha_{\text{зап}}^{\text{кул}}=\frac{2\pi e^2}{3nm^2c\omega}\left(\frac{2\mu}{\hbar^2}\right)^{5/2}\left|\frac{\partial}{\partial\boldsymbol{k}}[\boldsymbol{ep}_{cv}(0)]\right|^2\varepsilon_{\text{экс}}^{3/2}\left(1+\frac{\pi^2}{z^2}\right)\frac{e^z}{\operatorname{sh} z}. \qquad (2.34)$$

Далеко от края поглощения, когда $\hbar\omega-\varepsilon_0\gg\varepsilon_{\text{экс}}$, т. е. $z\ll 1$, выражение (2.34) переходит в (2.27). Вблизи края поглощения при $\hbar\omega\to\varepsilon_0$, т. е. $z\to\infty$, получим из (2.34)

$$\alpha_{\text{зап}}^{\text{кул}}\Big|_{\hbar\omega\to\varepsilon_0}=\frac{4\pi e^2}{3nm^2c\omega}\left(\frac{2\mu}{\hbar^2}\right)^{5/2}\left|\frac{\partial}{\partial\boldsymbol{k}}[\boldsymbol{ep}_{cv}(0)]\right|^2\varepsilon_{\text{экс}}^{3/2}. \qquad (2.35)$$

При поглощении фотона электрон и дырка могут перейти не только в состояние с непрерывным спектром, но и на дискретные уровни энергии связанных состояний экситона. Этим переходам будут соответствовать отдельные линии, расположенные непосредственно под краем поглощения. В упомянутой выше работе Эллиотт показал, что относительные интенсивности этих дискретных линий поглощения для разрешенных переходов пропорциональны $|\varphi_{\nu lm}(0)|^2$, где $\varphi_{\nu lm}(\boldsymbol{r})$ — волновая функция связанного состояния экситона, а ν, l, m — главное, азимутальное и магнитное квантовые числа этого состояния.

Из выражения для $\varphi_{\nu lm}(\boldsymbol{r})$ видно[1]), что при $\boldsymbol{r}=0$ отличны от нуля только s-состояния с $l=m=0$ и в этом случае

$$|\varphi_{\nu 00}(0)|^2=\frac{1}{\pi a_B^3}\frac{1}{\nu^3}, \qquad (2.36)$$

где главное квантовое число $\nu=1, 2, 3, \ldots$, а боровский радиус a_B равен (2.31б).

[1]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика, § 36, (36, 14).

Можно показать, что в этом случае коэффициент поглощения равен [1])

$$\alpha_{\text{раз}}^{\text{экс}} = \frac{8\pi}{cn}\frac{e^2 s^2}{\hbar a_B^3 \nu^3}\delta(\omega - \omega_\nu). \tag{2.37}$$

Здесь $s = \sqrt{\varepsilon_G/2\mu}$, ω_ν — частота, соответствующая дискретному терму экситона (2.30). Для запрещенных переходов в связанные экситонные состояния относительная интенсивность дискретных линий поглощения пропорциональна $|\partial\varphi_{\nu lm}(\boldsymbol{r})/\partial \boldsymbol{r}|^2_{r=0}$.

Можно показать, что для водородоподобного атома эта величина отлична от нуля только для p-состояний ($l = 1$). Представляя три нормированные функции p-состояния в виде $X = xf(r)$, $Y = yf(r)$ и $Z = zf(r)$, получим

$$\left|\frac{\partial\varphi_{\nu 1m}}{\partial \boldsymbol{r}}\right|^2_{r=0} = \left|\frac{\partial X(0)}{\partial x}\right|^2 + \left|\frac{\partial Y(0)}{\partial y}\right|^2 + \left|\frac{\partial Z(0)}{\partial z}\right|^2 = \frac{1}{3\pi a_B^5}\frac{\nu^2 - 1}{\nu^5}, \tag{2.38}$$

где главное квантовое число $\nu = 2, 3, \ldots$

Коэффициент поглощения в этом случае равен [2])

$$\alpha_{\text{зап}}^{\text{экс}} = \frac{16\pi}{3cn}\frac{e^2\hbar\theta^2}{m^{*2}a_B^3\omega}\frac{\nu^2 - 1}{\nu^5}\delta(\omega - \omega_\nu). \tag{2.39}$$

Здесь θ — безразмерная константа порядка единицы, а ω_ν — имеет тот же смысл, что и в формуле (2.37).

3. Вычисление матричного элемента $\mathcal{P}_{vc}$ (2.20) сопряжено со значительными трудностями, так как обычно нам неизвестны волновые функции электрона в валентной зоне и зоне проводимости. Однако во многих случаях уже достаточно определить, будет ли отличен от нуля интеграл

$$e\boldsymbol{M}_{cv} = e\int \psi_{ck}^*(\boldsymbol{r})\,\hat{\boldsymbol{p}}\,\psi_{vk}(\boldsymbol{r})\,d^3r, \tag{2.40}$$

которому пропорционален матричный элемент $\mathcal{P}_{vc} = \mathcal{P}_{12}$ (2.15).

Так как оператор $\hat{\boldsymbol{p}} = \frac{\hbar}{i}\frac{\partial}{\partial \boldsymbol{r}}$ преобразуется при операциях симметрии так же, как $\boldsymbol{r}$, то к интегралу (2.40) легко применить правила отбора (гл. II, § 10). Можно также показать, что интеграл (2.40) выражается через дипольный момент системы. В самом деле, оператор

$$\hat{\boldsymbol{p}} = \frac{1}{m}\hat{\boldsymbol{v}} = \frac{1}{m}\frac{\partial \boldsymbol{r}}{\partial t} = \frac{1}{m}\frac{1}{i\hbar}(\boldsymbol{r}\hat{\mathcal{H}} - \hat{\mathcal{H}}\boldsymbol{r}), \tag{2.41}$$

где мы воспользовались выражением для производной от оператора по времени [3]); здесь $\hat{\mathcal{H}}$ — гамильтониан системы.

---

[1]) Бир Г. Л., Пикус Г. Е. Симметрия и деформационные эффекты в полупроводниках.— М., 1972, с. 547.

[2]) Бир Г. Л., Пикус Г. Е., там же, с. 548.

[3]) Блохинцев Д. И., § 31.

Подставляя (2.41) в (2.40), получим

$$\boldsymbol{M}_{cv}=\frac{1}{im\hbar}\{(\hat{\boldsymbol{r}}\hat{\mathcal{H}})_{cv}-(\hat{\mathcal{H}}\boldsymbol{r})_{cv}\}=\frac{1}{im\hbar}\sum_{l}(\boldsymbol{r}_{cl}\mathcal{H}_{lv}-\mathcal{H}_{cl}\boldsymbol{r}_{lv})=$$

$$=\frac{1}{im\hbar}(\mathcal{E}_v\boldsymbol{r}_{cv}-\mathcal{E}_c\boldsymbol{r}_{cv})=\frac{\mathcal{E}_c-\mathcal{E}_v}{im\hbar}\langle c|\boldsymbol{r}|v\rangle. \qquad (2.42)$$

Мы воспользовались тем, что

$$\mathcal{H}_{lv}=\int\psi_l^*\hat{\mathcal{H}}\psi_v\,d^3r=\mathcal{E}_v\int\psi_l^*\psi_v\,d^3r=\mathcal{E}_v\delta_{lv} \qquad (2.42a)$$

и аналогично $\mathcal{H}_{cl}=\mathcal{E}_c\delta_{cl}$.

Таким образом,

$$e\boldsymbol{M}_{cv}=\text{const}\cdot\int\psi_{ck}^*r_e\psi_{vk}\,d^3r, \qquad (2.43)$$

где r_e — проекция радиуса-вектора на направление электрического поля $\boldsymbol{e}$.

Мы видим, что $\mathcal{P}_{vc}$ пропорционален матричному элементу дипольного перехода (2.43).

Правила отбора и, в частности, для дипольного перехода в поле симметрии $\boldsymbol{O}$, были рассмотрены в гл. II, § 10.

Рассмотрим правила отбора для дипольного перехода в центре бриллюэновской зоны ($\boldsymbol{k}=0$) для германия и кремния. Пусть переходы происходят из состояния Γ_{25}' в валентной зоне (см. рис. IV.27). В поле кубической симметрии $\boldsymbol{O}_h$ x, y, z принадлежат неприводимому представлению Γ_{15} (см. табл. III.2). Составляя по общему правилу (гл. II, § 10) прямое произведение $\Gamma_{15}\times\Gamma_{25}'$ и разлагая его по неприводимым представлениям группы $\boldsymbol{O}_h$, видим, что возможны только следующие переходы: $\Gamma_{25}'\leftrightarrow\Gamma_{15}$, $\Gamma_{25}'\leftrightarrow\Gamma_{25}$, $\Gamma_{25}'\leftrightarrow\Gamma_{12}'$, $\Gamma_{25}'\leftrightarrow\Gamma_2'$; остальные переходы из Γ_{25}' запрещены.

Более сложная ситуация возникает в теллуриде висмута, где правила отбора зависят от поляризации электромагнитной волны. Группа волнового вектора для точки $\boldsymbol{k}=0$ в Bi_2Te_3 — точечная группа $\boldsymbol{D}_{3d}=\boldsymbol{D}_3\times\boldsymbol{C}_i$; ее характеры даны в табл. VII.1. Легко установить, что координата z (параллельная оси кристалла c) преобразуется по представлению L_1', а координаты x, y — по представлению L_3'. Пусть электрическое поле волны направлено по z, тогда переход $L_1\leftrightarrow L_1'$ разрешен (в самом деле, прямое произведение $L_1'\times L_1$ содержит неприводимое представление L_1'). Если электрическое поле перпендикулярно z, то легко видеть, что переход $L_1\leftrightarrow L_1'$ запрещен, но разрешен переход $L_1\leftrightarrow L_3'$.

Если мы хотим учесть зависимость волновой функции от спина, необходимо рассматривать двойные группы и соответствующие им таблицы характеров. Процедура при этом усложняется, но в принципе остается такой же [1]).

[1]) Proc. of the Internat. School of Physics «Enrico Fermi». Course 34.— New York: Acad. Press, 1966, p. 93.

§ 3. Межзонные непрямые переходы

1. Для некоторых полупроводников (Ge, Si) энергетические зоны имеют вид, изображенный на рис. IV.28: самый глубокий минимум в зоне проводимости, расположен в точке, не совпадающей с положением максимума энергии в валентной зоне в точке O, при $\boldsymbol{k}=0$. В германии минимумы L в зоне проводимости расположены на границе зоны Бриллюэна в направлении ⟨111⟩, а в кремнии — на оси Δ в направлениях ⟨100⟩. В дальнейшем, для определенности, мы будем иметь в виду германий.

Переход электрона из валентной зоны (вблизи O) в зону проводимости (вблизи L) за счет поглощения кванта света невозможен. В самом деле, если $\boldsymbol{k}_1$ и $\boldsymbol{k}_2$ — волновые векторы электрона в начальном и конечном состояниях, а $\boldsymbol{\varkappa}$ — волновой вектор фотона, то из закона сохранения волнового вектора (квазиимпульса) следует: $\boldsymbol{k}_2-\boldsymbol{k}_1-\boldsymbol{\varkappa}\approx \boldsymbol{k}_2-\boldsymbol{k}_1\approx 0$, так как $\boldsymbol{\varkappa}\ll\boldsymbol{k}_2$, который порядка $\boldsymbol{k}_L$; учитывая, что k_1 мало́, мы видим, что закон сохранения волнового вектора выполняться не может. Если, однако, одновременно с поглощением фотона электрон поглотит (испустит) фонон с волновым вектором $\boldsymbol{q}=\pm(\boldsymbol{k}_2-\boldsymbol{k}_1)$, то закон сохранения волнового вектора выполняется.

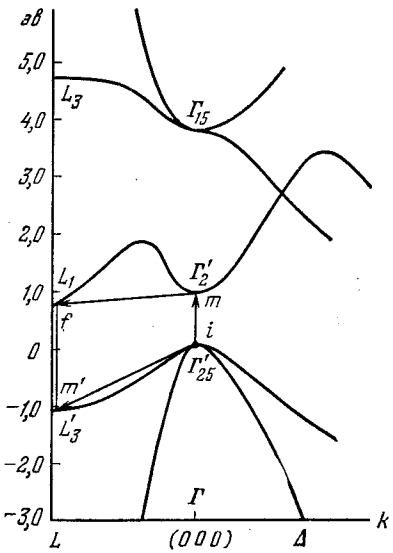


Рис. VII.1.

На рис. VII.1 представлены зоны германия и два канала imf и $im'f$ перехода электрона из начального состояния i в валентной зоне (вблизи $\boldsymbol{k}\approx 0$) в конечное состояние f в зону проводимости (вблизи точки L). На пути imf электрон переходит из начального состояния i в промежуточное m, поглощая фотон (прямой межзонный переход, рассмотренный в предыдущем параграфе), а затем переходит из промежуточного состояния m в конечное состояние f, поглощая (испуская) фонон с волновым вектором $\boldsymbol{q}=\pm(\boldsymbol{k}_2-\boldsymbol{k}_1)\approx\pm\boldsymbol{k}_L$. На пути $im'f$ электрон из валентной зоны с волновым вектором $\boldsymbol{k}_2\approx\boldsymbol{k}_L$ поглощает фотон и переходит в зону проводимости в окрестность точки L, а затем другой электрон валентной зоны вблизи точки O с волновым вектором $\boldsymbol{k}\approx 0$ поглощает (испускает) фонон $\boldsymbol{q}=\pm(\boldsymbol{k}_2-\boldsymbol{k}_1)$,

переходя на место дырки в точке $\boldsymbol{k}_2 \approx \boldsymbol{k}_L$. Последний процесс можно представить себе происходящим с дыркой: дырка в зоне проводимости поглощает фотон и переходит из состояния f в m' (это эквивалентно переходу электрона $m' \to f$), затем дырка поглощает фонон и переходит из состояния m' в i; оба перехода дырки эквивалентны переходу электрона из состояния i в f. Именно потому, что последний путь $im'f$ можно рассматривать и как переход электрона, и как переход дырки,— несущественно, занято ли состояние m' электроном или свободно.

Такие процессы с одновременным поглощением фотона и поглощением (испусканием) фонона должны считаться во втором приближении теории квантовых переходов (см. (П. 20.18)). Полное число таких переходов в единицу времени в единице объема через промежуточное состояние m равно

$$\mathscr{W} = \frac{2\pi}{\hbar} \sum_{if} \sum_{m} \frac{|M_{fm}^{\text{фон}}|^2 |M_{mi}^{\text{фот}}|^2}{(\varepsilon_m - \varepsilon_i - \hbar\omega)^2} \delta(\varepsilon_f - \varepsilon_i - \hbar\omega \mp \hbar\omega_q). \quad (3.1)$$

Здесь индексами i, f и m обозначены начальное, конечное и промежуточное состояния, $M_{mi}^{\text{фот}}$ — матричный элемент электронно-фотонного взаимодействия для перехода $i \to m$, $M_{fm}^{\text{фон}}$ — матричный элемент электронно-фононного взаимодействия для перехода $m \to f$, $\hbar\omega$ — энергия фотона, а $\hbar\omega_q$ — энергия фонона с волновым вектором $\boldsymbol{q}$.

Из теории квантовых переходов известно, что закон сохранения энергии должен выполняться только для начального состояния i и конечного f. Закон сохранения волнового вектора имеет место и для промежуточных переходов.

Мы можем с достаточным приближением описывать электроны и дырки в германии и кремнии в центре бриллюэновской зоны, т. е. вблизи точки $O(\boldsymbol{k}=0)$, скалярными эффективными массами m_c и m_v; тогда вблизи максимума энергии в валентной зоне для дырок

$$\varepsilon_i \equiv \varepsilon_v(\boldsymbol{k}) = \varepsilon_v(0) - \frac{\hbar^2 k^2}{2m_v} = \varepsilon_v(0) - \tilde{\varepsilon}_v, \quad (3.2)$$

и вблизи минимума энергии в зоне проводимости для электронов

$$\varepsilon_m \equiv \varepsilon_c(\boldsymbol{k}) = \varepsilon_c(0) + \frac{\hbar^2 k^2}{2m_c} = \varepsilon_c(0) + \tilde{\varepsilon}_c^0. \quad (3.3)$$

Здесь $\tilde{\varepsilon}_v$ и $\tilde{\varepsilon}_c^0$ — «кинетические» энергии дырок и электронов, пропорциональные k^2.

Поверхности постоянной энергии электронов в германии вблизи L, т. е. абсолютных минимумов, как мы знаем из гл. IV, § 15, — эллипсоиды вращения с поперечной эффективной массой m_t и продольной m_l; энергия электронов вблизи точки L

равна

$$\varepsilon_f \equiv \varepsilon_c(\boldsymbol{k}') = \varepsilon_c(\boldsymbol{k}_L) + \frac{\hbar^2}{2}\left(\frac{k_1^2+k_2^2}{m_t} + \frac{k_3^2}{m_l}\right) = \varepsilon_c(\boldsymbol{k}_L) + \tilde{\varepsilon}_c. \quad (3.4)$$

Здесь $\tilde{\varepsilon}_c$ — «кинетическая» энергия электрона вблизи точки L, зависящая, в главных осях, от квадратов прямоугольных составляющих волнового вектора.

Суммирование по начальному состоянию i, т. е. интегрирование по $\boldsymbol{k}$ (или $\tilde{\varepsilon}_v$), есть одновременно и суммирование по промежуточному состоянию m, так как при переходе $i \to m$ волновой вектор $\boldsymbol{k}$ сохраняется.

Используя (3.2)—(3.4), получим для разностей энергий, входящих в (3.1):

$$\varepsilon_m - \varepsilon_i - \hbar\omega = \varepsilon_c(\boldsymbol{k}) - \varepsilon_v(\boldsymbol{k}) - \hbar\omega = \varepsilon_0 + \tilde{\varepsilon}_c^0 + \tilde{\varepsilon}_v - \hbar\omega. \quad (3.5)$$

Здесь $\hbar\omega$ — энергия фотона и $\varepsilon_0 = \varepsilon_c(0) - \varepsilon_v(0)$ — ширина запрещенной зоны в точке O (для прямого перехода); далее

$$\begin{aligned}\varepsilon_f - \varepsilon_i - \hbar\omega \mp \hbar\omega_q &= \varepsilon_c(\boldsymbol{k}) - \varepsilon_v(\boldsymbol{k}) - \hbar\omega \mp \hbar\omega_q = \\ &= \varepsilon_c(\boldsymbol{k}_L) + \tilde{\varepsilon}_c - \varepsilon_v(0) + \tilde{\varepsilon}_v - \hbar\omega \mp \hbar\omega_q = \\ &= \varepsilon_G + \tilde{\varepsilon}_c + \tilde{\varepsilon}_v - \hbar\omega \mp \hbar\omega_q. \end{aligned} \quad (3.6)$$

Здесь $\mp\hbar\omega_q$ — энергии фонона, соответствующие его поглощению или испусканию, $\varepsilon_G = \varepsilon_c(\boldsymbol{k}_L) - \varepsilon_v(0)$ — ширина запрещенной зоны в Ge; $\varepsilon_0 - \varepsilon_G$ зависит от температуры и при 300°K равна $\varepsilon_0 - \varepsilon_G \approx (0{,}80-0{,}66)$ *эв* $= 0{,}14$ *эв*.

Для $M_{mi}^{\text{фот}}$ имеем, согласно (2.20),

$$M_{mi}^{\text{фот}} = A_0\left(\frac{e}{mc}\right)(\boldsymbol{e}\boldsymbol{p}_{cv}(\boldsymbol{k})) = \left(\frac{2\pi\hbar e^2 N}{m^2 c n \varkappa}\right)^{1/2}(\boldsymbol{e}\boldsymbol{p}_{cv}(\boldsymbol{k})), \quad (3.7)$$

где $\boldsymbol{p}_{cv}(\boldsymbol{k})$ равно (2.20а), а $|A_0|$ равно (2.12а); для разрешенных переходов $\boldsymbol{e}\boldsymbol{p}_{cv}(\boldsymbol{k}) \approx \boldsymbol{e}\boldsymbol{p}_{cv}(0)$ и не зависит от $\boldsymbol{k}$.

Исследуя в гл. VIII взаимодействие электрона с колебаниями решетки, мы увидим, что квадрат матричного элемента для поглощения и испускания фонона (VIII.3.10), (VIII.3.10а) содержит множитель N_q и $N_q + 1$, где N_q — функция Планка (III.11.10а):

$$N_q = \frac{1}{\exp\dfrac{\hbar\omega_q}{k_0 T} - 1}. \quad (3.8)$$

При температурах выше дебаевской ($k_0 T > \hbar\omega_q$)

$$N_q = N_q + 1 \approx \frac{k_0 T}{\hbar\omega_q}, \quad (3.8а)$$

при температурах $T \to 0$

$$N_q \to 0, \quad N_q + 1 \to 1, \quad (3.8б)$$

т. е. вероятность поглощения фонона стремится к нулю, но вероятность испускания фонона остается конечной.

Мы можем положить в самом общем случае

$$|M_{fm}^{\text{фон}}|^2 = C_q^{(j)}(N_q + {}^1/_2 \mp {}^1/_2). \tag{3.9}$$

Здесь j — номер ветви колебаний; верхний знак относится к случаю поглощения, нижний — к испусканию фонона. Выражение (3.9) можно считать не зависящим от $\boldsymbol{q}$. В самом деле, экспериментальные данные по колебательному спектру германия показывают, что энергия $\hbar\omega_q$ продольных акустических волн, взаимодействующих с электронами, меняется от нуля (для $\boldsymbol{q}=0$) до $2{,}75\cdot10^{-2}$ *эв* (для $\boldsymbol{q}\approx\boldsymbol{k}_L$); это составляет только 0,03 от величины $\varepsilon_G = 0{,}7$ *эв*, поэтому мы можем считать квадрат матричного элемента (3.9) и разность (3.6) не зависящими от $\boldsymbol{q}$ и положить, например, $\boldsymbol{q}=\boldsymbol{k}_L$ [1]).

Рассмотрим поглощение, обусловленное непрямыми разрешенными переходами, связанными с испусканием фононов j-й колебательной ветви.

Исходя из определения коэффициента поглощения света (2.22), получим

$$\alpha_{\text{раз}}^{\text{неп}} = 4\pi^2\frac{e^2}{m^2cn\omega}|\boldsymbol{e}\boldsymbol{p}_{cv}(0)|^2 C_{k_L}^{(j)}(N_{k_L j}+1)\times$$
$$\times\frac{2}{(2\pi)^3}\int d^3k\,\frac{2}{(2\pi)^3}\int d^3k'\,\frac{1}{(\varepsilon_m-\varepsilon_i-\hbar\omega)^2}\,\delta(\varepsilon_f-\varepsilon_i-\hbar\omega+\hbar\omega_q)+\ldots \tag{3.10}$$

Здесь знак плюс учитывает аналогичные члены, связанные со взаимодействием с другими ветвями колебаний j, а также с процессами поглощения фононов. Следует, однако, иметь в виду, что, как мы покажем в конце этого параграфа, существуют определенные правила отбора для матричных элементов $M_{fm}^{\text{фон}}$, так что не все фононы могут участвовать в непрямых переходах.

В (3.10) суммирование по начальному состоянию i и конечному f в (3.1) заменены интегрированием по $\boldsymbol{k}$ и $\boldsymbol{k}'$:

$$\sum_i \to \frac{2}{(2\pi)^3}\int d^3k = \frac{2}{(2\pi)^3}\cdot 2\pi\frac{(2m_v)^{3/2}}{\hbar^3}\int\sqrt{\tilde{\varepsilon}_v}\,d\tilde{\varepsilon}_v, \tag{3.10а}$$

$$\sum_f \to \frac{2}{(2\pi)^3}\int d^3k' = \frac{2}{(2\pi)^3}\cdot 2\pi\frac{(2m_{\text{эфф}})^{3/2}}{\hbar^3}\int\sqrt{\tilde{\varepsilon}_c}\,d\varepsilon_c, \tag{3.10б}$$

при этом мы воспользовались выражениями (VI.3.19а) и (VI.2.22); здесь $m_{\text{эфф}} = N_c^{2/3}(m_t^2 m_l)^{1/3}$.

---

[1]) На самом деле ситуация раз в десять более благоприятна, так как надо сравнить ε_G не с $\hbar\omega(\boldsymbol{k}_L) - \hbar\omega(0)$, а только с разбросом значений $\boldsymbol{q}$ при переходе к точке L.

Ввиду того, что поглощение, связанное с непрямыми переходами (второе приближение теории возмущений!), гораздо слабее поглощения, связанного с прямыми переходами, выражение (3.10) интересует нас только вблизи края непрямого поглощения: $\hbar\omega \approx \varepsilon_G$. Так как $\varepsilon_0 > \varepsilon_G$, а нас интересуют $\hbar\omega$, только немного превосходящие ε_G, то в (3.10) можно положить знаменатель

$$(\varepsilon_m - \varepsilon_i - \hbar\omega)^2 = (\varepsilon_0 + \tilde{\varepsilon}_c^0 + \tilde{\varepsilon}_v - \hbar\omega)^2 \approx (\varepsilon_0 - \hbar\omega)^2.$$

В результате (3.10) приобретает вид

$$\alpha_{\text{раз}}^{\text{неп}} = \frac{B_j^{\text{исп}}}{\omega(\varepsilon_0 - \hbar\omega)^2} \times$$

$$\times \int\int \sqrt{\tilde{\varepsilon}_v \tilde{\varepsilon}_c}\, \delta[\tilde{\varepsilon}_c + \tilde{\varepsilon}_v + \varepsilon_G + \hbar\omega_{k_L} - \hbar\omega]\, d\tilde{\varepsilon}_v\, d\tilde{\varepsilon}_c + \ldots, \qquad (3.11)$$

где $B_j^{\text{исп}}$ — константа, соответствующая испусканию фонона j-й ветви колебаний, не зависящая от ω, но зависящая от температуры. Используя в (3.11) для интегрирования по $\tilde{\varepsilon}_v$ δ-функцию, получим для двойного интеграла [1])

$$\int_0^b \sqrt{\tilde{\varepsilon}_c (b - \tilde{\varepsilon}_c)}\, d\tilde{\varepsilon}_c = \frac{\pi}{8} b^2, \qquad (3.11\text{а})$$

где

$$b = \hbar\omega - \varepsilon_G - \hbar\omega_{k_L}. \qquad (3.11\text{б})$$

Учитывая не только процессы испускания, но и поглощения фононов и суммируя по разным ветвям колебаний j, получим, используя (3.11а), (3.11б):

$$\alpha_{\text{раз}}^{\text{неп}} = \frac{B^{\text{исп}}}{\omega(\varepsilon_0 - \hbar\omega)^2}(\hbar\omega - \varepsilon_G - \hbar\omega_{k_L})^2 + \frac{B^{\text{пог}}}{\omega(\varepsilon_0 - \hbar\omega)^2}(\hbar\omega - \varepsilon_G + \hbar\omega_{k_L})^2. \qquad (3.12)$$

Для запрещенных переходов $\boldsymbol{ep}_{cv}(\boldsymbol{k}) = \frac{\partial}{\partial \boldsymbol{k}}[\boldsymbol{ep}_{cv}(\boldsymbol{k})]_{k=0}\boldsymbol{k}$, поэтому квадрат матричного элемента $M_{mi}^{\text{фот}}$ содержит лишний множитель k^2; наличие этого множителя под знаком интеграла в (3.10) приводит к тому, что коэффициент поглощения света $\alpha_{\text{зап}}^{\text{неп}}$ состоит из слагаемых, пропорциональных $(\hbar\omega - \varepsilon_G \mp \hbar\omega_{k_L})^3$.

2. Экситонные состояния влияют на коэффициент поглощения света $\alpha^{\text{неп}}$ не только при прямых, но и при непрямых переходах. Этот вопрос был исследован Р. Эллиоттом в упомянутой выше статье. Электронно-дырочное взаимодействие в случае непрямых переходов учитывается так же, как и при прямых, посредством весового множителя, дающего вероятность того, что

[1]) Бронштейн И. Н., Семендяев К. А., с. 364.

электрон и дырка находятся в одной точке пространства (§ 2, п. 2). Мы ограничимся приведением полученных результатов.

Из (3.12) видно, что без учета экситонных состояний коэффициент поглощения $\alpha_{\text{раз}}^{\text{неп}}$ для непрямых переходов состоит из суммы членов, пропорциональных $(\hbar\omega - \varepsilon_G \mp \hbar\omega_q)^2$. Учет взаимодействия электрона и дырки в случае перехода их при поглощении фотона к несвязанному состоянию изменяет в членах, определяющих $\alpha_{\text{раз}}^{\text{неп}}$, показатель степени 2 на 3/2. Если при непрямом переходе электрон и дырка образуют связанное (экситонное) состояние ниже уровня ε_G, то поглощение тоже имеет непрерывный характер и описывается степенными членами с показателем 1/2. В то же время мы для прямых переходов в этом случае имеем ряд дискретных линий. Непрерывный характер поглощения для непрямых переходов связан в том случае с тем, что в процессе участвует фонон, позволяющий переходить в любую точку экситонной зоны.

Наконец, показатель 3 для степенной зависимости $\alpha_{\text{зап}}^{\text{неп}}$ от энергии $\hbar\omega - \varepsilon_G \mp \hbar\omega_{k_L}$, в случае запрещенных непрямых переходов, переходит в 5/2 выше ε_G (несвязанные состояния) и в 3/2 ниже ε_G (связанные состояния).

3. Рассмотрим применение теории групп к выводу правил отбора для непрямых переходов в германии, в котором непрямые переходы были подробно изучены экспериментально. Известно (см. гл. IV, § 15), что зона проводимости германия состоит из восьми минимумов, центры которых расположены в точках L на границах зоны Бриллюэна (см. рис. IV.22).

На рис. IV.28, *а* схематически изображены энергетические зоны германия и указаны неприводимые представления, по которым преобразуется волновая функция электрона в симметричных точках зоны Бриллюэна. Теоретические и экспериментальные исследования показывают, что валентной зоне и зоне проводимости в точке $\boldsymbol{k} = 0$ соответствуют неприводимые представления Γ_{25}' и Γ_2' группы $\boldsymbol{O}_h$. Группа волнового вектора в точке L (см. рис. IV.22) является подгруппой группы $\boldsymbol{O}_h = \boldsymbol{T}_d \times \boldsymbol{C}_i$. Легко видеть, что вектор $\boldsymbol{k}_L$ переходит в самого себя при действии следующих элементов группы $\boldsymbol{T}_d$: E, $2C_3$, $3\sigma_v$, которые образуют группу $\boldsymbol{C}_{3v}$. Так как конец вектора $\boldsymbol{k}_L$ лежит на границе зоны Бриллюэна, то инверсия J переводит его в эквивалентный волновой вектор; таким образом, группой волнового вектора $\boldsymbol{k}_L$ является группа двенадцатого порядка $\boldsymbol{D}_{3d} = \boldsymbol{C}_{3v} \times \boldsymbol{C}_i$, состоящая из шести классов: E, $2C_3$, $3\sigma_v$, J, $2JC_3$, $3J\sigma_v$.

Мы представили группу $\boldsymbol{D}_{3d}$ волнового вектора $\boldsymbol{k}_L$ как прямое произведение: $\boldsymbol{D}_{3d} = \boldsymbol{C}_{3v} \times \boldsymbol{C}_i$; это, как мы увидим ниже, удобно, так как из гл. II, § 5, п. 2 нам известно, что элементы группы $\boldsymbol{T}_d \times \boldsymbol{C}_i$, а следовательно, и элементы ее подгруппы

$\boldsymbol{C}_{3v}\times\boldsymbol{C}_i$, связанные с инверсией J, только тогда являются операциями симметрии решетки германия, если они сопровождаются нетривиальной трансляцией $\boldsymbol{\alpha}=\frac{a}{4}(1, 1, 1)$. В табл. VII.1 представлены характеры группы $\boldsymbol{D}_{3d}$, при этом столбцам соответствуют одинаковые элементы симметрии, но один раз записанные через $\boldsymbol{D}_3\times\boldsymbol{C}_i$, а другой раз через $\boldsymbol{C}_{3v}\times\boldsymbol{C}_i$; в самом деле, $J\sigma_v=C_2'$ и, как отсюда следует, $\sigma_v=JC_2'$ (надо помнить, что $C_2'\perp C_3$ и σ_v проходит через ось C_3).

Таблица VII.1

$\boldsymbol{C}_{3v}\times\boldsymbol{C}_i$	E	$2C_3$	$3J\sigma_v$	J	$2JC_3$	$3\sigma_v$
$\boldsymbol{D}_3\times\boldsymbol{C}_i$	E	$2C_3$	$3C_2'$	J	$2JC_3$	$3JC_2'$
L_1	1	1	1	1	1	1
L_2	1	1	—1	1	1	—1
L_3	2	—1	0	2	—1	0
L_1'	1	1	1	—1	—1	—1
L_2'	1	1	—1	—1	—1	1
L_3'	2	—1	0	—2	1	0

Из расчета и эксперимента следует, что неприводимые представления электронной волновой функции в валентной зоне и зоне проводимости в точке L равны L_3' и L_1 (см. рис. VII.1).

Из рис. VII.1 мы видим, что в германии возможны непрямые переходы через ближайшие зоны по каналам: imf ($\Gamma_{25}'\to\Gamma_2'—L_1$) и $im'f$ ($\Gamma_{25}'\to L_3\to L_1$). В § 2, п. 3 мы показали, что прямой переход $\Gamma_{25}'\to\Gamma_2'$, связанный с поглощением фотона, разрешен. Так же просто показать, что разрешен прямой переход $L_3'\to L_1$. Таким образом, нам остается только рассмотреть правила отбора для переходов $\Gamma_2'\to L_1$ и $\Gamma_{25}'\to L_3'$, связанных с поглощением (испусканием) фонона с волновым вектором $\boldsymbol{q}_L$. Для этого предварительно определим, по каким неприводимым представлениям преобразуются нормальные колебания решетки германия в точке L. Мы должны действовать так, как в гл. III, § 8. Вначале мы должны определить полное приводимое представление, соответствующее всем колебательным степеням свободы в точке $\boldsymbol{q}_L$. Для германия число атомов в элементарной ячейке $s=2$, поэтому размерность этого полного представления равна $3s=3\cdot2=6$. Для определения характеров этого полного представления мы воспользуемся формулами (III.8.13)—(III.8.15).

Как было отмечено выше, элементы группы $\boldsymbol{D}_{3d}=\boldsymbol{C}_{3v}\times\boldsymbol{C}_i$ волнового вектора $\boldsymbol{q}_L$, связанные с инверсией J, должны сопро-

вождаться нетривиальной трансляцией $\boldsymbol{\alpha}=\frac{a}{4}(1,\ 1,\ 1)$. Так как при нетривиальной трансляции $\boldsymbol{\alpha}$ ни один из атомов элементарной ячейки не сохраняет своего положения ($n_{CJ}=0$), то из (III.8.15) следует, что

$$\chi^u(J\,|\,\boldsymbol{\alpha})=\chi^u(J\sigma_v\,|\,\boldsymbol{\alpha})=\chi^u(JC_3\,|\,\boldsymbol{\alpha})=0.$$

С другой стороны, при преобразованиях E и C_3, $n_{C(\Phi)}=2$, поэтому из (III.8.13) следует

$$\chi^u(E\,|\,0)=(1+2\cos 0°)\cdot 2=6,\quad \chi^u(C_3\,|\,0)=(1+2\cos 120°)\cdot 2=0.$$

Наконец, для преобразования σ_v из (III.8.14) получим ($n_{S(0)}=2$)

$$\chi^u(\sigma_v\,|\,0)=\chi^u(S(0)\,|\,0)=(-1+2\cos 0°)\cdot 2=2.$$

Объединяя эти результаты, выпишем характеры полного представления колебаний в точке L в табл. VII.2. Разлагая это приводимое представление по неприводимым представлениям табл. VII.1, получим

$$\chi^u=L_1+L_3+L_2'+L_3'. \tag{3.13}$$

Экспериментальное исследование рассеяния нейтронов в германии показало, что неприводимым представлениям нормальных колебаний (3.13) соответствуют:

$$\begin{aligned} &L_1 — LO \text{ — продольные оптические фононы,}\\ &L_3 — TA \text{ — поперечные акустические фононы,}\\ &L_2' — LA \text{ — продольные акустические фононы,}\\ &L_3' — TO \text{ — поперечные оптические фононы.} \end{aligned} \tag{3.14}$$

Перейдем теперь к правилам отбора для матричного элемента $M_{fm}^{\text{фон}}$, входящего в выражение (3.1). Рассматривая переход $\Gamma_2'\longrightarrow L_1$, отметим, что матричный элемент для этого перехода содержит под знаком интеграла электронные функции преобразующиеся по неприводимым представлениям Γ_2' и L_1, осцилляторные волновые функции нормальных колебаний кристалла и оператор $\hat{\mathcal{U}}$ взаимодействия электрона с фононом. Интегрирование по осцилляторным функциям, которое может быть произведено независимо от интегрирования по координатам электрона, приводит к правилам отбора (III.10.25), (III.10.26), согласно которым возможны только переходы, при которых только один фонон испускается или поглощается.

Таблица VII.2

$C_{3v}\times C_i$	E	$2C_3$	$3J\sigma_v$	J	$2JC_3$	$3\sigma_v$
χ^u	6	0	0	0	0	2

Можно показать, что оператор взаимодействия $\hat{\mathcal{U}}$ преобразуется по тем же неприводимым представлениям, что и нормальные координаты, т. е. по представлениям (3.13). Мы подойдем к вопросу несколько иначе. При разрешенном переходе, связанном с испусканием фонона, начальное состояние системы состоит из электрона, описываемого волновой функцией с симметрией Γ_2', конечное состояние — из электрона в состоянии L_1 и фонона, волновая функция которых должна обладать той же симметрией Γ_2', а это означает, что прямое произведение $\Gamma_2' \times L_1$ на одно из неприводимых представлений (3.13) должно содержать неприводимое представление Γ_2'; или, что то же самое, прямое произведение $\Gamma_2' \times L_1$ должно содержать одно из неприводимых представлений (3.13).

Характеры Γ_2', соответствующие классам группы $\boldsymbol{D_{3d}}$, можно взять из табл. III.2. Характеры прямого произведения $\Gamma_2' \times L_1$ будут равны значениям в табл. VII.3. Разлагая $\Gamma_2' \times L_1$ по неприводимым представлениям группы $\boldsymbol{D_{3d}}$, получим

$$\Gamma_2' \times L_1 = L_2', \tag{3.15}$$

как легко проверить по табл. VII.1.

Таблица VII.3

$\boldsymbol{D_{3d}}$	E	$2C_3$	$3C_2'$	J	$2JC_3$	$3JC_2'$
$\Gamma_2' \times L_1$	1	1	−1	−1	−1	1

Рассматривая второй канал, совершенно аналогично получим

$$\Gamma_{25}' \times L_3' = L_1' + L_2' + 2L_3'. \tag{3.16}$$

Сравнивая (3.15), (3.16) с (3.13) и (3.14), видим, что в непрямых переходах могут участвовать фононы:

$$\begin{aligned} L_2' &— LA — \text{продольный акустический фонон,} \\ 2L_3' &— TO — \text{поперечный оптический фонон.} \end{aligned} \tag{3.16}$$

При этом переход с участием LA-фонона может происходить через зону Γ_2' или L_3', а переход с участием TO-фонона происходит только через зону L_3'; множитель 2 при L_3' в (3.16) указывает, что имеется два линейно независимых матричных элемента, определяющих вероятность этого перехода. В то же время переходы с излучением или поглощением продольных оптических фононов и поперечных акустических фононов запрещены и в рассматриваемом нами процессе участвовать не могут, независимо от того, рассматривать ли переходы через зоны Γ_2' и L_3' или через

другие Γ и L; это следует из того, что произведение представлений Γ'_{25}, L_1 и представления Γ'_{15}, соответствующего поглощению фотона (см. § 2, п. 3), не содержат представлений L_3 и L'_2, связанных с поглощением фононов.

§ 4. Поглощение света в полупроводниках свободными носителями

При частотах, меньших частоты соответствующей краю межзонного поглощения света в полупроводниках, может наблюдаться слабое внутризонное поглощение свободными носителями.

Рассмотрим простую феноменологическую теорию поглощения света свободными носителями.

Мы будем исходить из уравнений Максвелла [1]) для однородной, изотропной, проводящей и немагнитной ($\mu = 1$) среды:

$$\operatorname{rot} \boldsymbol{H} = \frac{\varepsilon}{c}\dot{\boldsymbol{E}} + \frac{4\pi\sigma}{c}\boldsymbol{E}, \tag{4.1а}$$

$$\operatorname{rot} \boldsymbol{E} = -\frac{1}{c}\dot{\boldsymbol{H}}, \tag{4.1б}$$

$$\operatorname{div} \boldsymbol{E} = 0, \tag{4.1в}$$

$$\operatorname{div} \boldsymbol{H} = 0. \tag{4.1г}$$

Здесь $\boldsymbol{E}$ и $\boldsymbol{H}$ — напряженности электрического и магнитного полей, ε — диэлектрическая постоянная, определяющая вектор индукции $\boldsymbol{D} = \varepsilon\boldsymbol{E}$, σ — удельная электропроводность, определяющая ток $\boldsymbol{j}$ по закону Ома ($\boldsymbol{j} = \sigma\boldsymbol{E}$); свободные заряды в правой части (4.1в) отсутствуют.

Для того чтобы исключить $\boldsymbol{H}$ из (4.1а), (4.1б), возьмем ротор от обеих частей уравнения (4.1б) и производную по времени от обеих частей уравнения (4.1а); используя тождество [2]) $\operatorname{rot}\operatorname{rot} = \operatorname{grad}\operatorname{div} - \nabla^2$ и уравнение (4.1в), получим

$$\nabla^2\boldsymbol{E} = \frac{\varepsilon}{c^2}\ddot{\boldsymbol{E}} + \frac{4\pi\sigma}{c^2}\dot{\boldsymbol{E}}. \tag{4.2}$$

Для непроводящей среды ($\sigma = 0$) это уравнение превращается в волновое уравнение.

Уравнение (4.2) имеет частное решение в виде плоской волны с комплексным волновым вектором $\tilde{\boldsymbol{\varkappa}}$ (1.13); подставляя (1.13) в (4.2), получим

$$\tilde{\boldsymbol{\varkappa}}^2 = \frac{\omega^2}{c^2}\left(\varepsilon + \frac{4\pi\sigma}{\omega}i\right) = \frac{\omega^2}{c^2}\tilde{\varepsilon}. \tag{4.3}$$

---

[1]) Тамм И. Е., § 91.

[2]) Смирнов В. И., с. 365.

Отсюда комплексная диэлектрическая проницаемость

$$\tilde{\varepsilon} = \tilde{n}^2 = (n + ik)^2 = \varepsilon + \frac{4\pi\sigma}{\omega} i, \tag{4.4}$$

где $\tilde{n} = n + ik$ — комплексный показатель преломления.

При малых ω мнимая часть $\tilde{\varepsilon}$ равна $4\pi\sigma/\omega$, где σ — вещественная статическая электропроводность, не зависящая от частоты ω.

Покажем, что при больших частотах ω удельная электропроводность начинает зависеть от ω и становится комплексной. Как мы видели в § 1, п. 1, свободные носители, взаимодействующие только с электромагнитной волной, не могут поглощать света. Учтем взаимодействие свободных носителей с искажениями решетки (тепловыми колебаниями, примесями) посредством времени релаксации τ, так как это было сделано в гл. VI, § 6. Положим в (VI.6.4) магнитное поле $\boldsymbol{H} = 0$, тогда

$$v_x = \frac{eE_x}{m^*} \frac{\tau}{1 - i\omega\tau}. \tag{4.5}$$

Плотность тока вдоль оси x

$$j_x = \tilde{\sigma} E_x = N_0 e v_x = \frac{N_0 e^2 \tau}{m^*} \frac{E_x}{1 - i\omega\tau}, \tag{4.6}$$

где N_0 — концентрация свободных носителей; отсюда

$$\tilde{\sigma} = \sigma_0 \frac{1}{1 - i\omega\tau} = \sigma_0 \frac{1 + i\omega\tau}{1 + \omega^2\tau^2}, \tag{4.7}$$

где $\sigma_0 = N_0 e^2 \tau / m^*$ — значение удельной электропроводности в постоянном электрическом поле.

Подставляя (4.7) в (4.4), получим

$$\tilde{\varepsilon} = \varepsilon_1 + i\varepsilon_2 = (n + ik)^2 = \varepsilon - \frac{4\pi\sigma_0}{\omega} \frac{\omega\tau}{1 + \omega^2\tau^2} + i \frac{4\pi\sigma_0}{\omega} \frac{1}{1 + \omega^2\tau^2}. \tag{4.8}$$

Таким образом, вещественная и мнимая части диэлектрической проницаемости равны

$$\varepsilon_1 = n^2 - k^2 =$$

$$= \varepsilon - \frac{4\pi\sigma_0}{\omega} \frac{\omega\tau}{1 + \omega^2\tau^2} = \varepsilon - \frac{4\pi}{\omega} \operatorname{Im} \tilde{\sigma} = \varepsilon \left(1 - \frac{\omega_p^2}{\omega_0^2 + \omega^2}\right), \tag{4.9}$$

$$\varepsilon_2 = 2nk = \frac{4\pi\sigma_0}{\omega} \frac{1}{1 + \omega^2\tau^2} = \frac{4\pi}{\omega} \operatorname{Re} \tilde{\sigma} = \varepsilon \frac{\omega_0}{\omega} \frac{\omega_p^2}{\omega_0^2 + \omega^2}. \tag{4.9а}$$

Здесь $\omega_0 = 1/\tau$ — частота столкновений с нерегулярностями решетки, а $\omega_p = \sqrt{4\pi N_0 e^2 / \varepsilon m^*}$ — *плазменная частота*.

Из (4.9а) и (1.14) следует, что коэффициент поглощения

$$\alpha = \frac{2\omega k}{c} = \frac{4\pi}{nc} \operatorname{Re} \tilde{\sigma} = \frac{4\pi\sigma_0}{nc} \frac{1}{1 + \omega^2\tau^2}, \tag{4.9б}$$

т. е. α пропорционален $\mathrm{Re}\,\tilde{\sigma}$, что было уже использовано в гл. VI, § 6.

Из (4.9а), (4.9б) видно, что поглощение пропорционально σ_0, т. е. концентрации свободных носителей. В области $\omega^2\tau^2 \gg 1$ (но, конечно, $\omega < \omega_G$ — частота края межзонного поглощения) коэффициент поглощения $\alpha \sim \omega^{-2} \sim \lambda^2$, т. е. пропорционален квадрату длины волны; это подтверждается опытами над электронным германием (Х. Фэг и М. Беккер, 1951).

Развитая выше классическая теория поглощения света свободными носителями тока справедлива только в области $\hbar\omega < k_0T$ (или $\hbar\omega < \varepsilon_F$ — энергии Ферми, для вырожденных полупроводников). Последовательная квантовомеханическая теория поглощения света свободными носителями, основанная на теории возмущений во втором приближении, была развита Фрёлихом и последовательно изложена Фэном с сотрудниками. В предельном случае $\hbar\omega \gg k_0T$ для простой зоны и рассеяния на акустической ветви квантовая теория дает: $\alpha \sim \omega^{-2}(\hbar\omega/k_0T)^{1/2}$. Обзор теоретических результатов и сравнение с экспериментальными данными можно найти в статье Фэна [1]).

§ 5. Поляритоны

1. Рассмотрим некоторые новые элементарные возбуждения в твердом теле, которые являются своеобразным гибридом фононов оптической ветви колебаний кристалла с фотонами — квантами электромагнитного поля. Квазичастицы этих элементарных возбуждений получили название *поляритонов*.

В гл. III, § 9 мы исследовали полуфеноменологическим методом длинноволновые колебания оптической ветви ионного кубического кристалла, содержащего два разноименных иона в элементарной ячейке. Уравнения классического движения разноименных ионов друг относительно друга имеют вид (III.9.31)

$$\frac{d^2\boldsymbol{w}}{dt^2} = -\omega_0^2\boldsymbol{w} + \omega_0\sqrt{\frac{\varepsilon_0-\varepsilon_\infty}{4\pi}}\,\boldsymbol{E}. \tag{5.1}$$

Здесь $\boldsymbol{w} = \sqrt{N_0 m_r}\,\boldsymbol{s}$ — «нормированное» смещение ионов, где $\boldsymbol{s}$ — смещение положительного иона относительно отрицательного иона, N_0 — число ячеек в единице объема кристалла, m_r — приведенная масса ионов; $\omega_0^2 = \frac{\varkappa}{m_r} - \frac{4\pi N_0 e^{*}(\varepsilon_\infty+2)}{9m_r}$ — частота механических колебаний ионов, где $\varkappa$ — коэффициент квазиупругой силы взаимодействия ионов, e^* — эффективный заряд ионов, ε_0 и ε_∞ — статическая и высокочастотная диэлектрические постоянные, $\boldsymbol{E}$ — среднее электрическое поле в кристалле.

[1]) Fan H. — In: Semiconductors and Semimetals./Ed. by R. K. Willardson, A. C. Beer. — New York, London, 1967, v. 3, chap. 9.

Согласно (III.9.32) вектор поляризации

$$\boldsymbol{P}=\omega_0\sqrt{\frac{\varepsilon_0-\varepsilon_\infty}{4\pi}}\,\boldsymbol{w}+\frac{\varepsilon_\infty-1}{4\pi}\,\boldsymbol{E}. \tag{5.2}$$

Разделяя смещение $\boldsymbol{w}$ на поперечное $\boldsymbol{w}_t$ и продольные $\boldsymbol{w}_l$ ($\boldsymbol{w}=\boldsymbol{w}_t+\boldsymbol{w}_l$) (III.9.33), мы показали, что частоты поперечных и продольных колебаний равны (III.9.39), (III.9.40)

$$\omega_t=\omega_0,\quad \omega_l=\omega_0\sqrt{\frac{\varepsilon_0}{\varepsilon_\infty}}. \tag{5.3}$$

2. Для того чтобы построить теорию поляритонов, надо наряду с уравнением механических колебаний (5.1) использовать уравнения электромагнитного поля [1]):

$$\operatorname{div}\boldsymbol{D}=\operatorname{div}(\boldsymbol{E}+4\pi\boldsymbol{P})=0, \tag{5.4}$$

$$\operatorname{div}\boldsymbol{H}=0, \tag{5.5}$$

$$\operatorname{rot}\boldsymbol{E}=-\frac{1}{c}\dot{\boldsymbol{H}}, \tag{5.6}$$

$$\operatorname{rot}\boldsymbol{H}=\frac{1}{c}\dot{\boldsymbol{D}}=\frac{1}{c}(\dot{\boldsymbol{E}}+4\pi\dot{\boldsymbol{P}}). \tag{5.7}$$

Мы считаем среду немагнитной и предполагаем отсутствие свободных зарядов и токов проводимости.

В гл. III, § 9 мы воспользовались только одним из этих уравнений — (5.4), которое не учитывает запаздывания. Положим

$$\left.\begin{aligned}\boldsymbol{w}&=\boldsymbol{w}_0\\ \boldsymbol{P}&=\boldsymbol{P}_0\\ \boldsymbol{E}&=\boldsymbol{E}_0\\ \boldsymbol{H}&=\boldsymbol{H}_0\end{aligned}\right\}\times\exp i(\boldsymbol{kr}-\omega t), \tag{5.8}$$

где $\boldsymbol{w}_0$, $\boldsymbol{P}_0$, $\boldsymbol{E}_0$ и $\boldsymbol{H}_0$, вообще говоря, комплексные амплитуды.

Подставляя (5.8) в (5.1), (5.4)—(5.7), получим

$$-\omega^2\boldsymbol{w}=-\omega_0^2\boldsymbol{w}+\omega_0\sqrt{\frac{\varepsilon_0-\varepsilon_\infty}{4\pi}}\,\boldsymbol{E}, \tag{5.9}$$

$$\boldsymbol{k}(\boldsymbol{E}+4\pi\boldsymbol{P})=0, \tag{5.10}$$

$$\boldsymbol{kH}=0, \tag{5.11}$$

$$[\boldsymbol{kE}]=\frac{\omega}{c}\boldsymbol{H}, \tag{5.12}$$

$$[\boldsymbol{kH}]=-\frac{\omega}{c}(\boldsymbol{E}+4\pi\boldsymbol{P}). \tag{5.13}$$

В противоположность электростатическому случаю (гл. III, § 9), электрическое поле не может тождественно равняться нулю. В самом деле, если $\boldsymbol{E}=0$, то из (5.12) $\boldsymbol{H}=0$, но тогда из (5.13)

[1]) Тамм И. Е., § 91.

$\boldsymbol{P}=0$ и, наконец, из (5.2) $\boldsymbol{w}=0$; таким образом, при $\boldsymbol{E}=0$ реализуется тривиальный случай $\boldsymbol{E}=\boldsymbol{H}=\boldsymbol{P}=\boldsymbol{w}=0$.

Из (5.9) следует

$$\boldsymbol{w}=\frac{\omega_0}{\omega_0^2-\omega^2}\sqrt{\frac{\varepsilon_0-\varepsilon_\infty}{4\pi}}\,\boldsymbol{E}. \tag{5.14}$$

Подставляя это в (5.2), получим

$$\boldsymbol{P}=\left\{\frac{\omega_0^2}{\omega_0^2-\omega^2}\,\frac{\varepsilon_0-\varepsilon_\infty}{4\pi}+\frac{\varepsilon_\infty-1}{4\pi}\right\}\boldsymbol{E}. \tag{5.15}$$

Отсюда и из (5.10) получим

$$(\boldsymbol{kE})\left\{\frac{\omega_0^2}{\omega_0^2-\omega^2}\,(\varepsilon_0-\varepsilon_\infty)+\varepsilon_\infty\right\}=0. \tag{5.16}$$

Рассмотрим два случая. В первом фигурная скобка в (5.16) равна нулю, т. е.

$$\frac{\omega_0^2}{\omega_0^2-\omega^2}\,(\varepsilon_0-\varepsilon_\infty)+\varepsilon_\infty=0; \tag{5.17}$$

тогда из (5.15) следует $\boldsymbol{E}+4\pi\boldsymbol{P}=0$, но тогда из (5.13) следует $[\boldsymbol{kH}]=0$, что совместно с (5.11) дает $\boldsymbol{H}=0$, теперь из (5.12): $[\boldsymbol{kE}]=0$, но как мы раньше указывали, $\boldsymbol{E}\neq 0$, так что $\boldsymbol{E}\parallel\boldsymbol{k}$; используя (5.14) и (5.15), получим

$$\boldsymbol{w}\parallel\boldsymbol{P}\parallel\boldsymbol{E}\parallel\boldsymbol{k}, \tag{5.18}$$

т. е. все колеблющиеся векторы продольны, а частота, определяемая из (5.17), равна

$$\omega^2=\left(\frac{\varepsilon_0}{\varepsilon_\infty}\right)\omega_0^2=\omega_l^2. \tag{5.19}$$

Таким образом, это решение совпадает с электростатическим решением для продольных колебаний (III.9.42).

Во втором случае в (5.16) $\boldsymbol{kE}=0$, но так как $\boldsymbol{E}\neq 0$, то $\boldsymbol{E}\perp\boldsymbol{k}$; теперь из (5.12) следует, что $\boldsymbol{k}$, $\boldsymbol{E}$, $\boldsymbol{H}$ образуют в написанном здесь порядке правовинтовую ортогональную систему векторов, поэтому

$$kE=\frac{\omega}{c}H. \tag{5.20}$$

Колебания носят теперь поперечный характер, так как

$$\boldsymbol{H}\perp(\boldsymbol{E}\parallel\boldsymbol{P}\parallel\boldsymbol{w})\perp\boldsymbol{k} \tag{5.21}$$

и (5.11) выполняется автоматически. Подставляя в (5.13) (5.15) и (5.20), получим после сокращения на E

$$\frac{c^2k^2}{\omega^2}=\frac{\omega_0^2}{\omega_0^2-\omega^2}\,(\varepsilon_0-\varepsilon_\infty)+\varepsilon_\infty \tag{5.22}$$

(при этом мы учли, что $[\boldsymbol{kH}]$ направлено против E). Решая это

квадратное уравнение относительно ω^2, получим

$$\omega^2 = \frac{\varepsilon_0\omega_0^2 + c^2k^2 \pm \sqrt{(\varepsilon_0\omega_0^2 + c^2k^2)^2 - 4\varepsilon_\infty\omega_0^2c^2k^2}}{2\varepsilon_\infty}, \qquad (5.23)$$

т. е. две дисперсионные ветви.

Обозначим ветвь, соответствующую знаку плюс перед корнем через $\omega_1^2(k)$, а знаку минус перед корнем через $\omega_2^2(k)$. Тогда из (5.23) для $k \to 0$ получим

$$\omega_1(0) = \left(\frac{\varepsilon_0}{\varepsilon_\infty}\right)^{1/2}\omega_0 = \omega_t, \quad \omega_2(k \to 0) = \frac{c}{\sqrt{\varepsilon_0}}k, \qquad (5.24)$$

а для $k \to \infty$ получим

$$\omega_1(\infty) = \frac{c}{\sqrt{\varepsilon_\infty}}k, \quad \omega_2(\infty) = \omega_0. \qquad (5.24а)$$

При определении $\omega_2(\infty)$ следует учесть, что слагаемое c^2k^2 перед корнем сокращается с $(c^2k^2)^2$ под знаком корня, поэтому надо определить корень до следующего приближения по величине $(c^2k^2)^{-1}$.

На рис. VII.2 изображены зависимости $y = y(x)$, где $y = \omega/\omega_0$, а $x = \frac{k}{\omega_0/c} = ck/\omega_0$. Распространению света в вакууме соответствует пунктирная прямая $а$ с уравнением $y = x$ (или $\omega = ck$).

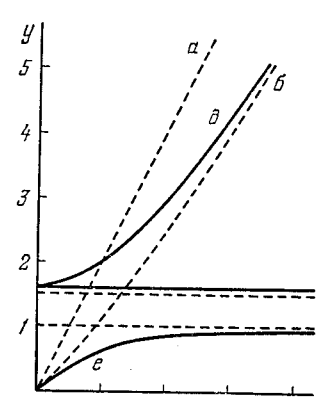


Рис. VII. 2.

В случае высоких частот $(k \to \infty)$ свет в среде распространяется по закону, которому соответствует уравнение $y = x/\sqrt{\varepsilon_\infty}$, т. е. $\omega = (c/\sqrt{\varepsilon_\infty})k$ (см. (5.24а)), а в случае низких частот $(k \to 0)$ — уравнение $y = x\sqrt{\varepsilon_0}$, т. е. $\omega = (c/\sqrt{\varepsilon_0})k$ (см. (5.24)); распространение света в среде изображается на рис. VII.2 пунктирной кривой $б$. Горизонтальные пунктирные прямые $в$ и $г$ изображают продольные и поперечные механические колебания с частотами $\omega_l = (\varepsilon_0/\varepsilon_\infty)^{1/2}\omega_0$ и $\omega_t = \omega_0$. Сплошные линии $в$, $д$, $е$ — дисперсионные кривые поляритонов, соответствующие решениям (5.19) и (5.23).

Сплошная прямая $в$ соответствует продольным поляритонам (5.18), (5.19) чисто механического характера. Ветви $д$ и $е$ описывают поперечные поляритоны смешанного электромагнитно-механического типа; они происходят от гибридизации колебаний, соответствующих пунктирным прямым $б$, $в$ и $г$. К. Хуанг [1]) вы-

[1]) Huang K.—Proc. Roy. Soc., 1951, v. A 208, p. 352.

числил долю механической энергии от общей энергии для поперечных поляритонов в ветвях $д$ и $е$. Он показал, что для тех же значений материальных констант, для которых вычислены графики на рис. VII.2, доля механической энергии в нижней ветви $е$ растет при увеличении $x=ck/\omega_0$, меняясь от 30% при $x=0$, до 95% при $x=6$; наоборот, для верхней ветви $д$ доля механической энергии поляритонов убывает при увеличении x, равняясь 70% при $x=0$ и уменьшаясь до 5% при $x=6$.

Поляритонная ветвь $д$ была исследована экспериментально.

Наряду с рассмотренными выше элементарными возбуждениями, являющимися гибридом фононов оптической ветви ионного кристалла и колебаний электромагнитного поля, существуют и другие гибриды квантов электромагнитного поля фотонов с возбуждениями твердого тела, например экситонами. Такие светоэкситоны, называемые сейчас тоже поляритонами, обнаружены на опыте. Их теория была существенно развита в работах С. И. Пекара.

§ 6. Эффект вращения Фарадея

1. В 1846 г. великий английский физик Майкл Фарадей обнаружил следующее явление. Если прозрачное тело поместить в сильное магнитное поле и пропустить плоскополяризованный луч света вдоль поля, то плоскость поляризации луча поворачивается на угол, пропорциональный напряженности магнитного поля и длине пути, пройденного лучом в магнитном поле.

Это явление, которое Фарадей, убежденный в единстве электромагнитных и световых явлений, искал в течение многих лет, стало важной вехой на пути становления электромагнитной теории света. Для полупроводников явление вращения Фарадея стало в ряде случаев удобным средством определения эффективной массы носителей тока.

Мы рассмотрим явление Фарадея в простейшем случае для свободных носителей тока со скалярной эффективной массой m^*.

В первую очередь мы должны немного обобщить теорию циклотронного резонанса, изложенную в гл. VI, § 6, п. 1. Рассмотрим не плоскополяризованную электромагнитную волну, а две поляризованные по кругу волны. В этом случае уравнение (VI.6.3а) приобретает вид

$$m^*\left(\dot{\boldsymbol{v}}+\frac{1}{\tau}\boldsymbol{v}\right)=-eE_0\boldsymbol{e}e^{-i\omega t}-\frac{e}{c}[\boldsymbol{v}\boldsymbol{H}]. \tag{6.1}$$

Здесь $\boldsymbol{v}$—скорость электрона, τ—время его релаксации, $\boldsymbol{H}$—напряженность постоянного магнитного поля, направленного вдоль оси z, E_0—амплитуда высокочастотного электрического

поля, $-e$ — заряд электрона и вектор поляризации $\boldsymbol{e}=\{1, \gamma, 0\}$ [1]), где $\gamma=\pm i=\cos\frac{\pi}{2}\pm i\sin\frac{\pi}{2}=\exp\left(\pm i\frac{\pi}{2}\right)$.

В (VI.6.3а) $\gamma=0$, поэтому электрическое поле имело только составляющую $E=E_0\exp(-i\omega t)$ вдоль оси x. В случае (6.1) электрическое поле наряду с составляющей $E_0\exp(-i\omega t)$ вдоль оси x имеет составляющую $E_0\exp\left[-i\left(\omega t\mp\frac{\pi}{2}\right)\right]$ вдоль оси y. Из элементарного курса оптики известно, что если колебания вдоль осей x и y имеют одинаковую амплитуду и отличаются по фазе на $\pm\pi/2$, то при сложении они дают колебание, поляризованное по кругу. При $\gamma=+i$ говорят о *левой круговой (циркулярной) поляризации* волны [2]), при $\gamma=-i$, — о *правой*.

Считая, что $\boldsymbol{v}$ тоже пропорциональна $\exp(-i\omega t)$, разложим уравнение (6.1) по осям x и y ($\boldsymbol{H}\parallel z$):

$$\begin{aligned}\frac{m^*}{\tau}(1-i\omega\tau)v_x&=-eE-\frac{e}{c}Hv_y,\\ \frac{m^*}{\tau}(1-i\omega\tau)v_y&=-eE\gamma+\frac{e}{c}Hv_x.\end{aligned}\tag{6.2}$$

Решая эти уравнения относительно v_x, получим для тока

$$j_x=\tilde{\sigma}E=-eN_0v_x=\sigma_0\frac{1-(i\omega+\gamma\omega_c)\tau}{(1-i\omega\tau)^2+\omega_c^2\tau^2}E,\tag{6.3}$$

где N_0 — концентрация свободных электронов, $\sigma_0=e^2N_0\tau/m^*$ — удельная электропроводность в постоянном электрическом поле и $\omega_c=eH/m^*c$ — циклотронная частота.

При $\gamma=0$ выражение (6.3) совпадает с (VI.6.6). Полагая в (6.3) $\gamma=\pm i$ и выделяя вещественную и мнимую части, получим для комплексной удельной электропроводности

$$\tilde{\sigma}=\sigma_{\mathrm{Re}}+i\sigma_{\mathrm{Im}},\tag{6.4}$$

где вещественная часть

$$\sigma_{\mathrm{Re}}=\sigma_0\frac{1+(\omega_c\pm\omega^2)\tau^2}{[1+(\omega_c^2-\omega^2)\tau^2]^2+4\omega^2\tau^2}\tag{6.4а}$$

и мнимая

$$\sigma_{\mathrm{Im}}=\sigma_0\frac{(\omega\pm\omega_c)(\omega_c^2-\omega^2)\tau^3-(\omega\mp\omega_c)\tau}{[1+(\omega_c^2-\omega^2)\tau^2]^2+4\omega^2\tau^2}.\tag{6.4б}$$

Как мы знаем из (4.9б), коэффициент поглощения α пропорционален σ_{Re}, т. е. вещественной части электропроводности. Если $\omega_c\tau\gg1$ (а именно в таких условиях проводят эксперимент по циклотронному резонансу), то из (6.4а) следует, что

---

[1]) Заметим, что $\boldsymbol{e}$ не единичный вектор.

[2]) При левой круговой поляризации электрический вектор в волне вращается по часовой стрелке, если смотреть вдоль волны.

в резонансе, когда $\omega \approx \omega_c$

$$\sigma_{\mathrm{Re}} \approx \begin{cases} \sigma_0 & \text{для } \gamma = +i, \quad (6.5а) \\ \frac{\sigma_0}{4\omega_c^2\tau^2} \approx 0 & \text{для } \gamma = -i, \quad (6.5б) \end{cases}$$

так как верхний знак в (6.4а) соответствует $\gamma = +i$, а нижний $-\gamma = -i$. Таким образом, для электронов наблюдается пик поглощения только для левой круговой поляризации. Так как дырки вращаются в направлении, противоположном электронам, то для них пик поглощения наблюдается только в случае $\gamma = -i$, т. е. для правой круговой поляризации.

Таким образом, зная характер круговой поляризации высокочастотного поля, можно определить, электроны или дырки участвуют в циклотронном резонансе.

2. Из (4.4) следует, что квадрат комплексного показателя преломления

$$\tilde{n}^2 = (n + ik)^2 = \varepsilon + i\frac{4\pi\tilde{\sigma}}{\omega} = \varepsilon + i\frac{4\pi}{\omega}(\sigma_{\mathrm{Re}} + i\sigma_{\mathrm{Im}}) =$$

$$= \varepsilon - \frac{4\pi}{\omega}\sigma_{\mathrm{Im}} + i\frac{4\pi}{\omega}\sigma_{\mathrm{Re}}, \quad (6.6)$$

откуда

$$n^2 - k^2 = \varepsilon - \frac{4\pi}{\omega}\sigma_{\mathrm{Im}}, \quad (6.6а)$$

$$2nk = \frac{4\pi}{\omega}\sigma_{\mathrm{Re}}. \quad (6.6б)$$

Если поглощение мало, так что им можно в (6.6а) пренебречь, то

$$n = \sqrt{\varepsilon - \frac{4\pi}{\omega}\sigma_{\mathrm{Im}}}. \quad (6.7)$$

Если мы имеем два луча, поляризованных по кругу в разных направлениях ($\gamma = \pm i$), то им соответствуют два разных значения $\sigma_{\mathrm{Im}}^{\pm}$ в (6.4б), а это значит, что показатели преломления этих лучей (6.7) тоже различны:

$$n^+ = \sqrt{\varepsilon - \frac{4\pi}{\omega}\sigma_{\mathrm{Im}}^+}, \quad (6.7а)$$

$$n^- = \sqrt{\varepsilon - \frac{4\pi}{\omega}\sigma_{\mathrm{Im}}^-}. \quad (6.7б)$$

Плоскополяризованная волна может быть представлена как наложение двух циркулярно-поляризованных в разных направлениях волн с одинаковыми амплитудами. Показатель преломления правополяризованной волны может отличаться от показателя преломления левополяризованной волны; в этом случае разность фаз между ними изменяется при их распространении в среде.

Обозначим фазы обеих циркулярно-поляризованных волн через

$$\varphi_{\pm} = \omega\left(\frac{n^{\pm} z}{c} - t\right). \tag{6.8}$$

Комплексная составляющая электрического поля вдоль оси x от обеих волн равна

$$E_x = E_0\left(e^{i\varphi_+} + e^{i\varphi_-}\right) = 2E_0 e^{i\frac{\varphi_+ + \varphi_-}{2}} \cos\frac{\varphi_+ - \varphi_-}{2}, \tag{6.9}$$

что легко проверить, если заменить косинус в правой части экспоненциальными функциями.

Так как фазы E_y в циркулярно-поляризованных волнах отличаются от $\varphi_{\pm}$ на $\pm\pi/2$, то

$$E_y = E_0\left[e^{i\left(\varphi_+ + \frac{\pi}{2}\right)} + e^{i\left(\varphi_- - \frac{\pi}{2}\right)}\right] = 2E_0 e^{i\frac{\varphi_+ + \varphi_-}{2}} \cos\left(\frac{\pi}{2} + \frac{\varphi_+ - \varphi_-}{2}\right). \tag{6.9а}$$

Если угол между плоскостью поляризации и осью x равен θ, то

$$\operatorname{tg}\theta = \frac{\operatorname{Re}\{E_y\}}{\operatorname{Re}\{E_x\}} = -\frac{\sin\frac{\varphi_+ - \varphi_-}{2}}{\cos\frac{\varphi_+ - \varphi_-}{2}} = -\operatorname{tg}\frac{\varphi_+ - \varphi_-}{2}, \tag{6.10}$$

где $\operatorname{Re}\{E_y\}$ и $\operatorname{Re}\{E_x\}$ — вещественные части E_y и E_x; в (6.10) было использовано то, что вещественные части одинаковых комплексных множителей в правой части (6.9) и (6.9а) совпадают. Абсолютная величина угла поворота плоскости поляризации при прохождении лучом света в магнитном поле расстояния $z = d$, равна

$$\theta = \frac{|\varphi_+ - \varphi_-|}{2} = \frac{\omega d}{2c}(n^+ - n^-), \tag{6.11}$$

как это следует из (6.10) и (6.8).

Из (6.7а), (6.7б) следует, что

$$(n^+)^2 - (n^-)^2 = \frac{4\pi}{\omega}\left(\sigma_{\mathrm{Im}}^- - \sigma_{\mathrm{Im}}^+\right)$$

или

$$n^+ - n^- = \frac{2\pi}{\omega\bar{n}}\left(\sigma_{\mathrm{Im}}^- - \sigma_{\mathrm{Im}}^+\right), \tag{6.12}$$

где $\bar{n} = {}^1/_2(n^+ + n^-)$.

Для определения эффективной массы m^* из эффекта Фарадея используют обычно область частот $\omega \gg \omega_c$, $1/\tau$; в этом случае из (6.4б) получим

$$\sigma_{\mathrm{Im}}^- - \sigma_{\mathrm{Im}}^+ = \sigma_0 \frac{2\omega_c}{\omega^2\tau^2}, \tag{6.13}$$

если подставить значения $\sigma_0 = e^2 N_0 \tau/m^*$ и $\omega_c = eH/m^*c$.

Из (6.11)—(6.13) получим для угла вращения плоскости поляризации

$$\theta = \frac{2\pi e^3 N_0 H d}{c^2 \bar{n} \omega^2 m^{*2}}. \tag{6.14}$$

В качестве $\bar{n}$ можно взять коэффициент преломления среды в отсутствие магнитного поля.

Существенно отметить, что выражение (6.14) не зависит от τ, т. е. механизма рассеяния; оно справедливо и в случае $\omega_c \tau \ll 1$, когда неприменим метод циклотронного резонанса. В частности, это позволяет измерять эффективную массу m^* при комнатных температурах. Формула (6.14) обобщается на случай сложной зонной структуры и носителей тока разных сортов; конечно, в этом случае она дает только некоторые средние значения эффективных масс.

Эффект Фарадея наблюдался и количественно сравнивался с результатами других измерений во многих полупроводниках: Ge, n-InSb, AlSb, GaP, GaAs, InAs и др. [1]).

§ 7. Теория межзонного поглощения света в квантующем магнитном поле

1. В § 2 мы рассмотрели поглощение света, связанное с прямыми межзонными переходами. Для прямых разрешенных переходов коэффициент поглощения $\alpha_{\text{раз}}$ определяется выражением (2.26).

Рассмотрим теперь теорию межзонных переходов в полупроводнике со стандартной зоной, помещенном в сильное (квантующее) магнитное поле.

В гл. VI, § 5, п. 3 мы рассмотрели движение электрона со скалярной эффективной массой m^* в квантующем магнитном поле (Л. Д. Ландау, 1930) и показали, что его волновая функция равна

$$F_{Nk_y k_z}(x, y, z) = e^{i(k_y y + k_z z)} \varphi_N(x - x_0). \tag{7.1}$$

Здесь $\varphi_N(x - x_0)$ — осцилляторная функция, отнесенная к положению равновесия $x_0 = -(\hbar c/eH) k_y$, N — квантовое число осциллятора. Для простоты мы предполагаем, что линейные размеры кристалла по всем трем направлениям равны единице. Собственные значения энергии, соответствующие собственной функции (7.1), равны

$$\varepsilon_{Nk_z} = (2N + 1)\mu^* H + \frac{\hbar^2 k_z^2}{2m^*}. \tag{7.2}$$

[1]) Обзор экспериментальных данных с подробным указанием литературы см. в книге: Optical properties of Solids./Ed. by F. Abelès.— Amsterdam, London: North-Holland Publ. Comp., 1972, p. 366.

Здесь $\mu^* = e\hbar/2m^*c$ — «эффективный» магнетон Бора; при этом $2\mu^* H = \hbar\omega_c$, где $\omega_c = eH/m^*c$ — циклотронная частота. Собственные значения энергии (7.2) вырождены по квантовому числу k_y.

Как показали И. Латтинджер и В. Кон (1955), волновая функция электрона в периодическом поле кристалла, помещенном в квантующее магнитное поле, в первом приближении, т. е. без учета взаимодействия с другими зонами, равна

$$\Psi(x, y, z) = u_{n0}(\boldsymbol{r})\, F_{Nk_yk_z}(x, y, z), \tag{7.3}$$

где $u_{n0}(\boldsymbol{r})$ — блоховская волновая функция электрона $u_{nk}(\boldsymbol{r})\, e^{i\boldsymbol{k}\boldsymbol{r}}$ в точке $\boldsymbol{k} = 0$ (предполагается простая невырожденная зона с минимумом энергии в точке $\boldsymbol{k} = 0$).

При наличии поля световой волны и внешнего магнитного поля гамильтониан $\hat{\mathcal{H}}$ (2.4) равен

$$\hat{\mathcal{H}} = \frac{1}{2m}\left[\hat{\boldsymbol{p}} + \frac{e}{c}\boldsymbol{A}^0 + \frac{e}{c}\boldsymbol{A}(\boldsymbol{r}, t)\right]^2 + V(\boldsymbol{r}), \tag{7.4}$$

где для постоянного магнитного поля, направленного вдоль оси z, вектор-потенциал $\boldsymbol{A}^0 = \{0, Hx, 0\}$ (см. (VI.5.11)).

Возмущением в (7.4), как и в (2.4), является поле электромагнитной волны $\boldsymbol{A}(\boldsymbol{r}, t)$ (2.8). Поступая аналогично (2.5), получим вместо (2.7) возмущение

$$\hat{\mathcal{H}}' = \frac{e}{mc}\boldsymbol{A}(\boldsymbol{r}, t)\left[\hat{\boldsymbol{p}} + \frac{e}{c}\boldsymbol{A}^0\right]. \tag{7.5}$$

Матричный элемент этого возмущения между волновыми функциями (7.3) для валентной зоны и зоны проводимости равен

$$\mathcal{P}_{vc} = \langle \Psi_c | \hat{\mathcal{H}}' | \Psi_v \rangle = \int u^*_{c0} F^*_{\nu'} \hat{\mathcal{H}}' u_{v0} F_\nu \, d^3r. \tag{7.6}$$

Здесь Ψ_c и Ψ_v — волновые функции (7.3) для зоны проводимости и валентной зоны, $\nu \equiv \{N, k_y, k_z\}$ — квантовые числа электрона в квантующем магнитном поле.

Подынтегральное выражение (7.6) содержит быстро меняющиеся в пределах ячейки кристалла Ω_0 функции $u^*_{c0}(\boldsymbol{r})$ и $u_{v0}(\boldsymbol{r})$ и медленно меняющиеся множители $F_{\nu'}(\boldsymbol{r})$ и $F_\nu(\boldsymbol{r})$. Перейдем в (7.6) от интегрирования по основной области кристалла $V = Z\Omega_0$ ($V = 1$ см^3) к сумме интегралов по объемам кристаллических ячеек Ω_0:

$$\begin{aligned}
\mathcal{P}_{vc} &= \int_V u^*_{c0} F^*_{\nu'} \frac{e}{mc}\boldsymbol{A}\left(\hat{\boldsymbol{p}} + \frac{e}{c}\boldsymbol{A}^0\right) u_{v0} F_\nu \, d^3r = \\
&= \frac{e}{mc}\int_V u^*_{c0} F^*_{\nu'}\left[(\boldsymbol{A}\hat{\boldsymbol{p}} u_{v0})\, F_\nu + (\boldsymbol{A}\hat{\boldsymbol{p}} F_\nu)\, u_{v0} + \frac{e}{c}\boldsymbol{A}\boldsymbol{A}^0 u_{v0} F_\nu\right] d^3r = \\
&= \frac{e}{mc}\sum_z \left\{ F^*_{\nu'} F_\nu \int_{\Omega_0} u^*_{c0}(\boldsymbol{A}\hat{\boldsymbol{p}} u_{v0})\, d^3r + \right. \\
&\left. + (F^*_{\nu'}\boldsymbol{A}\hat{\boldsymbol{p}} F_\nu)\int_{\Omega_0} u^*_{c0} u_{v0}\, d^3r + \frac{e}{c}\boldsymbol{A}\boldsymbol{A}^0 F^*_{\nu'} F_\nu \int_{\Omega^0} u^*_{c0} u_{v0}\, d^3r \right\}.
\end{aligned} \tag{7.7}$$

При этом мы вынесли из под знаков интегралов по элементарным ячейкам Ω_0 медленно меняющиеся функции.

Из-за ортогональности функции u_{c0} и u_{v0} два последних слагаемых в фигурной скобке (7.7) равны нулю и

$$\mathcal{P}_{vc} = \sum_z F^*_{\nu'} F_\nu \Omega_0 \cdot \frac{1}{\Omega_0} \int_{\Omega_0} u^*_{c0} \frac{e}{mc} \hat{\boldsymbol{A}\boldsymbol{p}} u_{v0}\, d^3r = \int_V F^*_{\nu'} F_\nu\, d^3r \cdot \frac{1}{\Omega_0} \int_{\Omega_0} u^*_{c0} \frac{e}{mc} \hat{\boldsymbol{A}\boldsymbol{p}} u_{v0}\, d^3r. \qquad (7.8)$$

где мы заменим суммирование по Z интегралом по объему V.

Полагая $\boldsymbol{A}$ равным (2.8) $p = -i\hbar\nabla$ и полагая, как и в § 2, волновой вектор фотона $\varkappa \approx 0$, получим, используя (7.1),

$$\mathcal{P}_{vc} = \frac{e}{mc} A_0 (\boldsymbol{e}\boldsymbol{p}_{cv}) \int e^{i(k_y - k'_y)y}\, dy \int e^{i(k_z - k'_z)z} dz \times \int \varphi^*_{N'}(x - x'_0)\, \varphi_N(x - x_0)\, dx, \qquad (7.9)$$

где $\boldsymbol{p}_{cv}$ равно (2.20а).

Интеграл по области $L_y = 1$ *см* равен

$$\int_0^1 e^{i(k_y - k'_y)y} dy = \left\{ \begin{array}{l} 1, \text{ если } k_y = k'_y, \\ 0, \text{ если } k_y \neq k'_y \end{array} \right\} = \delta_{k_y k'_y},$$

т. е. символу Кронекера; аналогично для интеграла по dz. Так как осцилляторные функции $\varphi_N(x - x_0)$ ортонормированы, а $x'_0 = x_0$ (из-за того, что $k_y = k'_y$), то последний интеграл в (7.9) равен $\delta_{NN'}$. В результате вместо (7.9) получим

$$\mathcal{P}_{vc} = \frac{e}{mc} A_0 (\boldsymbol{e}\boldsymbol{p}_{cv}) \delta_{k_y k'_y} \delta_{k_z k'_z} \delta_{NN'}. \qquad (7.9а)$$

Для вычисления числа переходов в единице объема в единицу времени, под действием возмущения $\hat{\mathcal{H}}'$ (7.5), необходимо суммировать (интегрировать) при вычислении $\mathcal{W}_{vc}$ как по начальному, так и конечному состояниям, т. е. по квантовым числам $\nu \equiv \{N, k_y, k_z\}$ и $\nu' \equiv \{N', k'_y, k'_z\}$. В силу (7.9а) это двойное суммирование может быть сведено к суммированию по N, k_y, k_z.

Таким образом, аналогично (2.21) получим

$$\mathcal{W}_{vc} = \frac{2\pi}{\hbar} \sum_{N, k_y, k_z} |\mathcal{P}_{vc}|^2 \delta(\varepsilon_c - \varepsilon_v - \hbar\omega) = \frac{2\pi}{\hbar} \frac{e^2}{m^2 c^2} |A_0|^2 |\boldsymbol{e}\boldsymbol{p}_{cv}(0)|^2 \sum_N \int \frac{2dk_y}{2\pi} \int \frac{2dk_z}{2\pi} \delta \left[\hbar\omega - \varepsilon_G - (2N+1)\mu_r H - \frac{\hbar^2 k_z^2}{2m_r} \right], \qquad (7.10)$$

где обратная приведенная эффективная масса $1/m_r = 1/m_c + 1/m_v$

и «приведенный эффективный магнетон Бора» $\mu_r = e\hbar/2m_rc$; в аргументе δ-функции мы положим $N' = N$ (так как $k_y' = k_y$) и $k_z' = k_z$.

Поскольку $k_y = -\frac{eH}{\hbar c}x_0$, а $0 \leqslant x_0 \leqslant 1$, интегрирование по k_y дает

$$\int dk_y = \frac{eH}{\hbar c}. \tag{7.11}$$

Далее

$$\int dk_z \delta\left[B - \frac{\hbar^2 k_z^2}{2m_r}\right] = \frac{1}{\hbar}\left(\frac{m_r}{2}\right)^{1/2} B^{-1/2}, \tag{7.12}$$

в чем легко убедиться, заменив переменную интегрирования k_z на $\zeta = \hbar^2 k_z^2/2m_r$.

Коэффициент поглощения света α_H равен числу переходов $\mathscr{W}_{vc}$ (7.10), деленному на поток фотонов $Nv = Nc/n$ [1]); если подставить в (7.10) вместо A_0 (2.12а) и воспользоваться (7.11) и (7.12), то мы получим

$$\alpha_H = \frac{2e^2}{m^2\omega nc}|\boldsymbol{ep}_{cv}(0)|^2\left(\frac{2m_r}{\hbar^2}\right)^{3/2}\mu_r H \sum_N [\hbar\omega - \varepsilon_G - (2N+1)\mu_r H]^{-1/2}. \tag{7.13}$$

Сумма по N в (7.13) распространяется на все значения N, для которых подрадикальные выражения не отрицательны. Те значения ω, N и H, для которых подрадикальное выражение в (7.13) равно нулю, определяют сингулярные точки коэффициента поглощения α_H. Эти точки соответствуют условию

$$\hbar\omega_{\max} - \varepsilon_G = (2N+1)\mu_r H, \tag{7.14}$$

где $\omega_{\max}$ — частота света, соответствующая максимуму поглощения. Конечно, на самом деле коэффициент поглощения α_H не обращается в этих точках в бесконечность, так как существует ряд факторов (например, взаимодействие электронов с колебаниями решетки), которые «замывают» бесконечные нити. Такие конечные пики (осцилляции) α_H наблюдаются на опыте.

Рис. VII.3.

Из (7.14) видно, что при заданном N частота света $\omega_{\max}$ линейно зависит от магнитного поля H. На рис. VII.3 представлены зависимости частоты света $\omega_{\max}$ от магнитного поля H для разных N. Вертикальная прямая $H = \text{const}$ пересекает полупря-

[1]) Здесь N — число фотонов в 1 $см^3$.

мые для разных N, определяя при заданном H соответствующие пикам поглощения значения ω_{max}. Точки пересечения горизонтальной прямой $\omega = \text{const}$ определяют при данной частоте ω_{max} значения магнитного поля H, соответствующие пикам поглощения.

Наблюдаемая на опыте картина часто более сложна: кривая поглощения состоит не только из пиков, но и из ступенек; это связано с тем, что наряду с прямыми разрешенными переходами, рассмотренными выше, могут существовать непрямые переходы с участием фононов и запрещенные переходы, которые ведут себя по-разному при $\boldsymbol{e} \parallel \boldsymbol{H}$ и $\boldsymbol{e} \perp \boldsymbol{H}$ ($\boldsymbol{e}$ — вектор поляризации в (2.8)).

Рассмотрим кратко непрямые переходы. При непрямых переходах мы должны в (3.1) отдельно суммировать по N и N' и интегрировать по k_y и k'_y, k_z и k'_z (в этом случае равенство (7.9а) не имеет места, так как во взаимодействии участвует фонон со своим волновым вектором $\boldsymbol{q}$). Двойное интегрирование по k_y и k'_y дает, согласно (7.11), множитель $(eH/\hbar c)^2$; опуская постоянный множитель, получим

$$\alpha^H_{\text{неп}} \sim \left(\frac{eH}{\hbar c}\right)^2 \sum_{N\,N'} \iint dk_z\, dk'_z \delta \left[(2N'+1)\,\mu_c H + \frac{\hbar^2 k'^2_z}{2mc} + \right.$$
$$\left. + \varepsilon_G + (2N+1)\,\mu_v H + \frac{\hbar^2 k_z^2}{2m_v} \mp \hbar\omega_q - \hbar\omega\right]. \qquad (7.15)$$

В аргументе δ-функции появляется дополнительное, по сравнению с (7.10), слагаемое $\mp \hbar\omega_q$, связанное с поглощением (верхний знак) или испусканием (нижний знак) фонона.

Введем переменные: $(\hbar^2 k_z^2/2m_v) = x^2$, $(\hbar^2 k'^2_z/2m_c) = y^2$ и $x^2 + y^2 = r^2$, тогда двойной интеграл в (7.15) будет пропорционален величине

$$\int_0^\infty d(r^2)\,\delta\left[r^2 - (\hbar\omega \pm \hbar\omega_q - \varepsilon_{Gi}) - (2N+1)\,\mu_v H - (2N'+1)\,\mu_c H\right], \qquad (7.15а)$$

где мы воспользовались тем, что $dk_z\, dk'_z \sim dx\, dy \sim r\, dr = {}^1/_2 d(r^2)$.

Из свойств δ-функции известно, что

$$\int_0^\infty \delta(x-a)\,dx = \theta(a) = \begin{cases} 0, & \text{если } a < 0, \\ 1, & \text{если } a > 0. \end{cases} \qquad (7.16)$$

«Ступенчатая» θ-функция используется в математике наряду с δ-функцией.

В результате мы можем для (7.15) написать

$$\alpha^H_{\text{неп}} \sim \left(\frac{eH}{\hbar c}\right)^2 \sum_{N\,N'} \theta\left[\hbar\omega \pm \hbar\omega_q - \varepsilon_G - (2N+1)\,\mu_v H - (2N'+1)\,\mu_c H\right]. \qquad (7.17)$$

Рассмотрим зависимость $\alpha^H_{\text{неп}}$ от энергии фотона $\hbar\omega$ при постоянном магнитном поле. Выберем слагаемое в сумме (7.17) с $N = N' = 0$. Для малых $\hbar\omega$, пока аргумент функции $\theta(u)$ меньше нуля $\theta(u) = 0$; пусть при увеличении $\hbar\omega$ в точке *1* на рис. VII.4 аргумент функции $\theta(u)$ становится равным нулю, тогда для бóльших значений $\hbar\omega$ функция $\theta(u) = 1$ и $\alpha^H_{\text{неп}}$ испытывает скачок, изображенный на рисунке. В точке *2* для другого слагаемого, например $N = 1$, $N' = 0$, аргумент функции $\theta(u)$ становится равным нулю, так что в $\alpha^H_{\text{неп}}$ появляется вторая ступенька. Из-за того, что $\mu_v \neq \mu_c$ и энергия фонона $\hbar\omega_q$ разная для разных колебательных ветвей и может входить со знаком плюс или минус в аргумент функции $\theta(u)$, ступеньки в $\alpha^H_{\text{неп}}$ имеют нерегулярный характер, как это и изображено на рис. VII.4. Прямые запрещенные переходы, которые мы не будем рассматривать, в еще большей мере усложняют картину поглощения света в полупроводниках в магнитном поле.

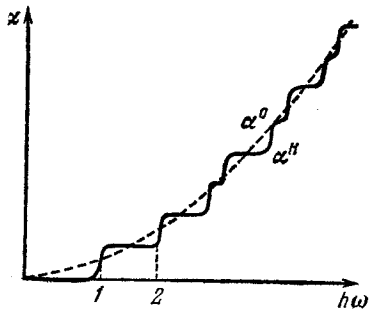


Рис. VII. 4.

2. Новая полезная информация о полупроводниках может быть получена при изучении поглощения света в скрещенных электрическом и магнитном полях (А. Г. Аронов, 1963).

Пусть магнитное поле $\boldsymbol{H}$ направлено по оси z, а электрическое поле $\boldsymbol{E}$ — по оси x. Выберем вектор-потенциал, как и в гл. VI, § 5, п. 3 в виде $\boldsymbol{A} \equiv \{0, Hx, 0\}$.

Уравнение Шредингера для электрона в зоне проводимости с эффективной массой m_c и зарядом $-e$ имеет вид

$$\hat{\mathcal{H}} F(\boldsymbol{r}) = \varepsilon_c F(\boldsymbol{r}), \tag{7.18}$$

где гамильтониан

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m_c}\frac{\partial^2}{\partial x^2} + \frac{1}{2m_c}\left(\frac{\hbar}{i}\frac{\partial}{\partial y} + \frac{e}{c}Hx\right)^2 - \frac{\hbar^2}{2m_c}\frac{\partial^2}{\partial z^2} + eEx. \tag{7.18а}$$

Здесь потенциальная энергия электрона в электрическом поле равна $-(-e)Ex = eEx$.

Так как гамильтониан $\hat{\mathcal{H}}$ не содержит явно y и z, то решение (7.18), как и в гл. VI, § 5, п. 3, ищем в виде

$$F(x, y, z) = \varphi(x)\, e^{i(k_y y + k_z z)}. \tag{7.19}$$

Действуя на (7.19) гамильтонианом $\hat{\mathcal{H}}$ и сокращая на экспонен-

циальный множитель, получим

$$-\frac{\hbar^2}{2m_c}\frac{d^2\varphi(x)}{dx^2}+\left[\frac{1}{2m_c}\left(\frac{eH}{c}\right)^2x^2+\left(\frac{\hbar eH}{m_cc}k_y+eE\right)x+\frac{\hbar^2(k_y^2+k_z^2)}{2m_c}\right]\varphi(x). \tag{7.20}$$

Это выражение может быть приведено к виду

$$-\frac{\hbar^2}{2m_c}\frac{d^2\varphi(x)}{dx^2}+\frac{1}{2}m_c\omega_c^2(x-x_c)^2\varphi(x)+w_c\varphi(x), \tag{7.21}$$

где ω_c, x_c и w_c — постоянные, не зависящие от x. Сравнивая коэффициенты при x^2, x^1, $x^0=1$ в (7.20) и (7.21), получим

$$\omega_c=\frac{eH}{m_cc}, \tag{7.21а}$$

$$x_c=-\lambda^2k_y-\frac{eE\lambda^2}{\hbar\omega_c}, \tag{7.21б}$$

$$w_c=\frac{\hbar^2k_z^2}{2m_c}-\lambda^2eEk_y-\frac{m_cc^2}{2}\left(\frac{E}{H}\right)^2. \tag{7.21в}$$

Здесь ω_c — циклотронная частота для электрона проводимости с массой m_c, $\lambda=(\hbar/m_c\omega_c)^{1/2}=(\hbar c/eH)^{1/2}$ — *магнитная длина*. Уравнение Шредингера (7.18) для функции $\varphi(x)$ приобретает вид

$$-\frac{\hbar^2}{2m_c}\frac{d^2\varphi(x)}{dx^2}+\frac{1}{2}m_c\omega_c^2(x-x_c)^2\varphi(x)=(\varepsilon_c-w_c)\varphi(x), \tag{7.22}$$

аналогичный (VI.5.18).

Таким образом, собственные значения энергии

$$\varepsilon_c=\varepsilon(0)+(2N+1)\mu_cH+w_c=$$
$$=\varepsilon_c(0)+(2N+1)\mu_cH+\frac{\hbar^2k_z^2}{2m_c}-\lambda^2eEk_y-\frac{m_cc^2}{2}\left(\frac{E}{H}\right)^2, \tag{7.23}$$

где $\varepsilon_c(0)$ — уровень энергии, соответствующий нижнему краю зоны проводимости $\mu_c=e\hbar/2m_cc$ — эффективный магнетон Бора, соответствующий электрону, $N=0, 1, 2, \ldots$ — квантовое число осциллятора. Мы видим, что в скрещенных электрическом и магнитном полях собственные значения энергии ε_c зависят от k_y, т. е. вырождение по k_y, которое имело место в магнитном поле, снимается.

Собственные функции уравнения (7.22) равны

$$\varphi(x)=\varphi_N(x-x_c)=\frac{\exp\left[-\frac{1}{2}\left(\frac{x-x_c}{\lambda}\right)^2\right]}{\sqrt{\lambda}}H_N\left(\frac{x-x_c}{\lambda}\right), \tag{7.24}$$

т. е. имеют тот же вид (VI.5.19а), что и при наличии одного магнитного поля (конечно, с другой постоянной x_c (7.21б)); H_N — полиномы Эрмита.

Для того чтобы определить коэффициент поглощения света в скрещенных электрическом и магнитном полях $\alpha^{H/E}$ для прямых разрешенных переходов необходимо, в соответствии с (7.9), знать волновые функции (7.19) и собственные значения энергии (7.23) и для дырок в валентной зоне.

Для дырок в валентной зоне с зарядом $+e$ и эффективной массой m_v получим вместо (7.21а) — (7.21в)

$$\omega_v = \frac{eH}{m_v c}, \tag{7.25а}$$

$$x_v = -\lambda^2 k_y' + \frac{eE\lambda^2}{\hbar\omega_v}, \tag{7.25б}$$

$$w_v = \frac{\hbar^2 k_z'^2}{2m_v} + \lambda^2 eEk_y' - \frac{m_v c^2}{2}\left(\frac{E}{H}\right)^2, \tag{7.25в}$$

где мы заменили $-e$ на e и m_c на m_v. Для собственного значения энергии дырки вместо (7.23) получим

$$\varepsilon_v = \varepsilon_v(0) - (2N'+1)\mu_v H - w_v = \\ = \varepsilon_v(0) - (2N'+1)\mu_v H - \frac{\hbar^2 k_z'^2}{2m_v} - \lambda^2 eEk_y' + \frac{m_v c^2}{2}\left(\frac{E}{H}\right)^2. \tag{7.26}$$

Здесь $\mu_v = e\hbar/2m_v c$ — эффективный магнетон Бора, соответствующий дырке. При этом мы энергию дырки отсчитываем от края валентной зоны $\varepsilon_v(0)$ в отрицательном направлении («вниз»). Можно было бы вместо этого рассматривать состояние электрона в валентной зоне с отрицательной эффективной массой $-m_v$, при этом мы получили бы для ε_v то же выражение (7.26). Матричный элемент $\mathcal{P}_{vc}$ (7.9) будет теперь равен

$$\mathcal{P}_{vc} = \frac{fe}{mc} A_0 (\boldsymbol{e}\boldsymbol{p}_{cv}(0)) \delta_{k_y k_y'} \delta_{k_z k_z'} \int_{-\infty}^{\infty} \varphi_N(x - x_c)\varphi_{N'}(x - x_v)\, dx. \tag{7.27}$$

Мы видим, что как и в случае (7.9), $k_y = k_y'$, $k_z = k_z'$, однако теперь $x_c \neq x_v$ (так как $m_c \neq m_v$); из-за того, что центры тяжести осцилляторных функций теперь не совпадают, они не ортогональны друг к другу, т. е. в общем случае интеграл отличен от нуля при $N \neq N'$.

Для его вычисления можно воспользоваться формулой [1])

$$\int_{-\infty}^{\infty} e^{-x^2} H_m(x+y) H_n(x+z)\, dx = 2^n \sqrt{\pi}\, m!\, z^{n-m} \mathcal{L}_m^{n-m}(-2yz), \tag{7.28}$$

где $m \leqslant n$, а $\mathcal{L}_m^{n-m}$ — обобщенные полиномы Лагерра [2]). Преобра-

[1]) Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений.—М.: Наука, 1962, с. 852, (7.377).

[2]) Лебедев Н. Н. Специальные функции и их применение. —М.—Л., 1963, с. 100.

зуем интеграл в (7.27), положив $(x-x_c)/\lambda=\xi-a$ и выберем a так, чтобы исчез линейный по ξ член в показателе экспоненты.

В этом случае интеграл в (7.27) равен

$$\exp\left[-\left(\frac{x_c-x_v}{2\lambda}\right)^2\right]\int\limits_{-\infty}^{\infty} e^{-\xi^2}H_N\left(\xi-\frac{x_c-x_v}{2\lambda}\right)H_{N'}\left(\xi+\frac{x_c-x_v}{2\lambda}\right)d\xi=$$

$$=\exp\left[-\left(\frac{x_c-x_v}{2\lambda}\right)^2\right]\sqrt{\pi}\,2^N N'!\left(\frac{x_v-x_c}{2\lambda}\right)^{N-N'}\mathscr{L}_{N'}^{N-N'}\left[2\left(\frac{x_c-x_v}{2\lambda}\right)^2\right], \tag{7.29}$$

где $N'\leqslant N$ (в случае $N'>N$ надо поменять местами N с N' и x_v с x_c); $\mathscr{L}_{N'}^{N-N'}(u)$ — полиномы степени N' от u; первые из них равны

$$\mathscr{L}_0^p(u)=1,\quad \mathscr{L}_1^p(u)=1+p-u,$$
$$\mathscr{L}_2^p(u)={}^1/_2[(1+p)(2+p)-2(2+p)u+u^2].$$

Мы можем теперь $\mathscr{P}_{vc}$ (7.27) записать в виде

$$\mathscr{P}_{vc}=\frac{e}{mc}A_0(\boldsymbol{ep}_{cv}(0))\,\delta_{k_y k_y'}\delta_{k_z k_z'}\sqrt{\pi}\,e^{-\gamma^2}2^N N'!\,\gamma^{N-N'}\mathscr{L}_{N'}^{N-N'}(2\gamma^2), \tag{7.30}$$

где

$$\gamma=\frac{x_v-x_c}{2\lambda}=\frac{(m_c+m_v)\,c\lambda E}{2\hbar H}. \tag{7.30а}$$

Аргумент δ-функции, входящей в $\mathscr{W}_{vc}$ (7.10), равен

$$\varepsilon_c-\varepsilon_v-\hbar\omega=\varepsilon_G+(2N+1)\,\mu_c H+(2N'+1)\,\mu_v H+$$
$$+\frac{\hbar^2k_z^2}{2m_r}-\frac{(m_c+m_v)\,c^2}{2}\left(\frac{E}{H}\right)^2-\hbar\omega, \tag{7.31}$$

где мы использовали то, что $k_y=k_y'$ и $k_z=k_z'$. Поступая далее, как и в предыдущем пункте, получим для коэффициента поглощения в скрещенных электрическом и магнитном полях для прямых разрешенных переходов

$$\alpha^{H/E}=\frac{2e^2}{m^2\omega nc}|\boldsymbol{ep}_{cv}(0)|^2\left(\frac{2m_r}{\hbar^2}\right)^{3/2}(\mu_r H)\,e^{-2\gamma^2}\sum_{NN'}\pi 2^{2N}[N'!]^2\,\gamma^{2(N-N')}\times$$
$$\times[\mathscr{L}_{N'}^{N-N'}(2\gamma^2)]^2\left[\hbar\omega-\varepsilon_G-(2N+1)\,\mu_c H-(2N'+1)\,\mu_v H+\right.$$
$$\left.+\frac{(m_c+m_v)\,c^2}{2}\frac{E^2}{H^2}\right]^{-1/2}, \tag{7.32}$$

где $N'\leqslant N$.

Отметим основные особенности поглощения света в скрещенных электрическом и магнитном полях:

а) не существуют правила отбора по квантовым числам магнитных осцилляторов N и N', т. е. в принципе переход может осуществляться между любыми подуровнями Ландау в валентной зоне и зоне проводимости;

б) с ростом электрического поля $\alpha^{H/E}$ убывает экспоненциально; при больших полях $\alpha^{H/E} \sim E^{2(N'+N)}e^{-\beta E^2}$. Это налагает ограничение на величину максимального электрического поля, при котором вообще может наблюдаться поглощение: $E_{\max} \approx \approx \hbar H/(m_c+m_v)c\lambda$ ($\gamma^2 \sim 1$);

в) для разрешенных при $\boldsymbol{E}=0$ переходов ($N=N'$) смещение пика поглощения в электрическом поле равно: $\Delta\hbar\omega=(m_c+m_v)\times \times c^2E^2/2H^2$; измеряя это смещение, можно определить сумму эффективных масс m_c+m_v. Так как из измерения α_H (7.14) можно определить относительную эффективную массу $m_r= =m_cm_v/(m_c+m_v)$, то оба измерения $\alpha^{H/E}$ и α_H позволяют определить эффективные массы m_c и m_v в отдельности; эти же массы можно определить из частот переходов между уровнями с $N \neq N'$.

Эффективные массы m_c и m_v могут быть определены и по разности частот переходов, запрещенных при $\boldsymbol{E}=0$, но возгорающихся в электрическом поле.

§ 8. Поглощение света в полупроводниках в однородном электрическом поле (эффект Франца — Келдыша)

1. В 1958 г. независимо В. Франц и Л. В. Келдыш теоретически исследовали поглощение света в полупроводнике в однородном электрическом поле.

Задача может быть решена по общей схеме изложенной в предыдущем параграфе. Мы будем рассматривать полупроводник с простой зоной и будем считать, что электрон в зоне проводимости описывается скалярной эффективной массой m_c, а в валентной зоне — отрицательной эффективной массой $-m_v$.

В качестве функции $F_{Nk_yk_z}$ в (7.3) мы должны взять стационарную волновую функцию свободного электрона в однородном электрическом поле $\boldsymbol{E}$.

Как будет видно из дальнейшего, квантовомеханическую задачу о движении свободного электрона в однородном электрическом поле проще решать не в x-представлении, а в p-представлении. Это просто означает, что в уравнении Шредингера надо считать операторами не составляющие импульса ($\hat{p}_x=-i\hbar\partial/\partial x$ и т. д.), а координаты ($\hat{x}=i\hbar\partial/\partial p_x$ и т. д.); в этом случае стационарная волновая функция F зависит не от x, y, z, а от p_x, p_y, p_z[1]). Если электрическое поле $\boldsymbol{E}$ направлено вдоль оси x, то потенциальная энергия электрона с зарядом $-e$ равна $-(-e)Ex=eEx$. Уравнение Шредингера для электрона в зоне проводимости с массой m_c в поле $\boldsymbol{E}\parallel x$ в p-представлении

[1]) Блохинцев Д. И., § 13.

имеет вид

$$\left[\frac{p_x^2+p_y^2+p_z^2}{2m_c}+eEi\hbar\frac{\partial}{\partial p_x}\right]F_c(p_x,\ p_y,\ p_z)=\varepsilon_c F_c(p_x,\ p_y,\ p_z). \quad (8.1)$$

Здесь $F_c(p_x,\ p_y,\ p_c)$—волновая функция электрона в p-представлении, ε_c—собственное значение его энергии.

Из (8.1) видно, что в p-представлении волновая функция удовлетворяет дифференциальному уравнению первого порядка, в то время как в x-представлении она удовлетворяет дифференциальному уравнению второго порядка [1]), что делает в математическом отношении задачу более сложной.

Полагая

$$F_c(p_x,\ p_y,\ p_z)=\psi_c(p_x)\,\varphi_c(p_y)\,\chi_c(p_z), \quad (8.2)$$

подставляя это выражение в (8.1) и деля обе части уравнения на $F_c(p_x,\ p_y,\ p_z)$, получим

$$\left(\frac{p_x^2}{2m_c}+\frac{1}{\psi_c(p_x)}eEi\hbar\frac{d\psi_c}{dp_x}\right)+\frac{p_y^2}{2m_c}+\frac{p_z^2}{2m_c}=\varepsilon_c. \quad (8.3)$$

Первое слагаемое (в круглых скобках) зависит только от p_x, второе—только от p_y, третье—только от p_z. Так как сумма их постоянна и равна ε_c, то каждое слагаемое тоже постоянно; для того чтобы удовлетворить этому, положим $p_y=p_{yc}$ и $p_z=p_{zc}$, где p_{yc} и p_{zc}—собственные значения составляющих импульса по y и z; тогда вместо (8.3) получим

$$\frac{d\psi_c}{dp_x}-\frac{i}{\hbar eE}\left(\frac{p_x^2}{2mc}-\varepsilon_c+\frac{1}{2m_c}p_{\perp c}^2\right)\psi_c=0, \quad (8.4)$$

где $p_{\perp c}^2=p_{yc}^2+p_{zc}^2$.

Волновая функция (8.2) равна

$$F_c(p_x,\ p_y,\ p_z)=\psi_c(p_x)\,\delta_{p_y p_{yc}}\delta_{p_z p_{zc}}, \quad (8.5)$$

где $\delta_{p_y p_{yc}}$, $\delta_{p_z p_{zc}}$—символы Кронекера (мы считаем кристалл ограниченным, так что p_y и p_z принимают квазидискретные значения).

Решение линейного дифференциального уравнения первого порядка (8.4) имеет вид [2])

$$\psi_c(p_x)=C\exp\left\{\frac{i}{\hbar eE}\int_0^{p_x}\left(\frac{p_x'^2}{2m_c}-\varepsilon_c+\frac{1}{2m_c}p_{\perp c}^2\right)dp_x'\right\}=$$

$$=C\exp\left\{\frac{i}{\hbar eE}\left[\frac{p_x^3}{6m_c}-\left(\varepsilon_c-\frac{1}{2m_c}p_{\perp c}^2\right)p_x\right]\right\}. \quad (8.6)$$

Коэффициент C определим из условия нормировки волновой

[1]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика, 1973, § 24.
[2]) Смирнов В. И. т. II, с. 23.

функции $\psi_c(p_x)$ на δ-функцию по энергии, т. е.

$$\int_{-\infty}^{\infty} \psi_c^*(p_x,\ \varepsilon_c')\,\psi_c(p_x,\ \varepsilon_c)\,dp_x = \delta(\varepsilon_c - \varepsilon_c'). \tag{8.7}$$

Подставляя сюда в левую часть (8.6), получим

$$|C|^2 \int_{-\infty}^{\infty} \exp\left\{\frac{i}{\hbar eE}(\varepsilon_c - \varepsilon_c')\,p_x\right\} dp_x =$$

$$= |C|^2 (\hbar eE) \int_{-\infty}^{\infty} \exp\{i(\varepsilon_c - \varepsilon_c')\,\xi\}\,d\xi = |C|^2 (2\pi\hbar eE)\,\delta(\varepsilon_c - \varepsilon_c').$$

Здесь мы воспользовались стандартным определением δ-функции [1]). Сравнивая последнее выражение с (8.7), видим, что

$$C = 1/\sqrt{2\pi\hbar eE}. \tag{8.8}$$

Для электрона в валентной зоне надо решить уравнение аналогичное (8.3), заменив в нем m_c на $-m_v$ и ε_c на $\varepsilon_v + \varepsilon_G$ (ε_G — ширина запрещенной зоны); последнее для того, чтобы привести отсчет энергии в валентной зоне и зоне проводимости к одному уровню; тогда

$$F_v(p_x,\ p_y,\ p_z) = \psi_v(p_x)\,\delta_{p_y p_{yv}}\delta_{p_z p_{zv}}, \tag{8.9}$$

где

$$\psi_v(p_x) = C\exp\left\{-\frac{i}{\hbar eE}\left[\frac{p_x^3}{6m_v} + \left(\varepsilon_v + \varepsilon_G + \frac{1}{2m_v}p_{\perp v}^2\right)p_x\right]\right\}. \tag{8.10}$$

Нормировочная константа C в (8.10) имеет то же значение, что и в (8.8). Определим теперь матричный элемент $\langle F_c | F_v\rangle$, входящий в $\mathcal{P}_{vc}$ (7.8); имеем

$$\langle F_c | F_v\rangle = \sum_{p_y p_z} \delta_{p_y p_{yc}}\delta_{p_z p_{zc}}\delta_{p_y p_{yv}}\delta_{p_z p_{zv}} \times$$

$$\times |C|^2 \int_{-\infty}^{\infty} dp_x \exp\left\{-\frac{i}{\hbar eE}\left[\frac{p_x^3}{6\mu} + \left(\varepsilon_v - \varepsilon_c + \varepsilon_G + \frac{1}{2m_c}p_{\perp c}^2 + \right.\right.\right.$$

$$\left.\left.\left. + \frac{1}{2m_v}p_{\perp v}^2\right)p_x\right]\right\}. \tag{8.11}$$

Здесь μ приведенная масса:

$$\frac{1}{\mu} = \frac{1}{m_c} + \frac{1}{m_v}. \tag{8.11а}$$

Суммирование по p_y и p_z дает $\delta_{p_{yc}p_{yv}}\delta_{p_{zc}p_{zv}} = \delta_{p_{\perp c}p_{\perp v}}$. Заменим переменную интегрирования p_x, положив

$$p_x = (2\mu\hbar eE)^{1/3}\,u. \tag{8.12}$$

[1]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика, 1973, (15,7).

Тогда вместо (8.11) получим

$$\langle F_c | F_v \rangle = \frac{\delta_{p_{\perp c} p_{\perp v}}}{\hbar\omega_E} \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \exp\left[-i\left(\frac{u^3}{3}+xu\right)\right] du, \qquad (8.13)$$

где частота

$$\omega_E = (eE)^{2/3}/(2\mu\hbar)^{1/3}, \qquad (8.13а)$$

а

$$x = \frac{1}{\hbar\omega_E}\left[\varepsilon_v - \varepsilon_c + \varepsilon_G + \frac{1}{2m_c} p_{\perp c}^2 + \frac{1}{2m_v} p_{\perp v}^2\right]. \qquad (8.13б)$$

Так как

$$\int\limits_{-\infty}^{\infty} \exp\left[-i\left(\frac{u^3}{3}+xu\right)\right] du = \int\limits_{0}^{\infty} \left[\exp i\left(\frac{u^3}{3}+xu\right) + \right.$$

$$\left. + \exp\left[-i\left(\frac{u^3}{3}+xu\right)\right]\right] du = 2\int\limits_{0}^{\infty} \cos\left(\frac{u^3}{3}+xu\right) du,$$

то

$$\langle F_c | F_v \rangle = \frac{\delta_{p_{\perp v} p_{\perp c}}}{\hbar\omega_E} Ai(x), \qquad (8.14)$$

где *функция Эйри*[1])

$$Ai(x) = \frac{1}{\pi}\int\limits_{0}^{\infty} \cos\left(\frac{u^3}{3}+xu\right) du. \qquad (8.14а)$$

Аналогично (7.8)

$$\mathscr{P}_{vc} = \frac{e}{mc} A_0 (\boldsymbol{e}\boldsymbol{p}_{cv}(0)) \langle F_c | F_v \rangle = \frac{e}{mc} A_0 (\boldsymbol{e}\boldsymbol{p}_{cv}(0)) \delta_{p_{\perp v} p_{\perp c}} \frac{1}{\hbar\omega_E} Ai(x), \qquad (8.15)$$

где мы использовали (8.14). Наличие символа Кронекера $\delta_{p_{\perp v} p_{\perp c}}$ позволяет при вычислении $\mathscr{W}_{vc}$ (7.10) суммировать (интегрировать) только по $p_{\perp c} = p_{\perp v} \equiv p_\perp$, ε_v и ε_c; таким образом, коэффициент поглощения света в электрическом поле

$$\alpha_E = \frac{\mathscr{W}_{vc}}{Nc/n} = \frac{2\pi}{\hbar}\frac{n}{Nc} \sum_{\varepsilon_v,\, \varepsilon_c,\, p_\perp} |\mathscr{P}_{vc}|^2 \delta(\varepsilon_c - \varepsilon_v - \hbar\omega) =$$

$$= \frac{2\pi}{\hbar}\frac{n}{Nc}\frac{e^2}{m^2c^2} |A_0|^2 |\boldsymbol{e}\boldsymbol{p}_{cv}(0)|^2 \frac{1}{(\hbar\omega_E)^2} \times$$

$$\times \int d\varepsilon_v \int d\varepsilon_c\, 2\int\limits_{0}^{\infty} \frac{2\pi p_\perp dp_\perp}{(2\pi\hbar)^2} Ai^2(x)\, \delta(\varepsilon_c - \varepsilon_v - \hbar\omega), \qquad (8.16)$$

где множитель 2 перед последним интегралом учитывает спин электрона.

---

[1]) Джеффрис Г., Свирлс Б. Методы математической физики.—М.: Мир, 1970, вып. 3, с. 59.

При интегрировании по ε_v мы полагаем в x $\varepsilon_v = \varepsilon_c - \hbar\omega$, тогда из (8.13б)

$$x = \frac{1}{\hbar\omega_E}\frac{1}{2\mu}p_\perp^2 + \frac{\omega_G - \omega}{\omega_E}. \tag{8.16а}$$

Для электрического поля $E = 3000$ *в/см* и $2\mu = 10^{-27}$ *г*, частота $\omega_E = 3\cdot 10^{12}$ *сек*$^{-1}$. В (8.16а) $\omega_G = \varepsilon_G/\hbar$ — частота, соответствующая ширине запрещенной зоны; для $\varepsilon_G \approx 1$ *эв* $\omega_G \approx 10^{15}$ *сек*$^{-1}$, т. е. ω_G на два-три порядка больше ω_E.

Если длина образца в направлении x равна $L_x = 1$ *см*, то при фиксированном $p_\perp$ энергия ε_c меняется от $\varepsilon_c(0)$ до $\varepsilon_c(L_x) = \varepsilon_c(0) + eEL_x = \varepsilon_c(0) + eE$, поэтому интегрирование по ε_c дает множитель

$$\int\limits_{\varepsilon_c(0)}^{\varepsilon_c(L_x)} d\varepsilon_c = eE. \tag{8.16б}$$

Переходя посредством (8.16а) от переменной интегрирования $p_\perp$ к x и подставляя вместо A_0 выражение (2.12а), получим из (8.16)

$$\alpha_E = R(\hbar\omega_E)^{1/2}\pi\int\limits_\beta^\infty Ai^2(x)\,dx, \tag{8.17}$$

где

$$R = \frac{2e^2}{m^2 cn\omega}|\boldsymbol{ep}_{cv}(0)|^2\left(\frac{2\mu}{\hbar^2}\right)^{3/2} \tag{8.17а}$$

и

$$\beta = \frac{\omega_G - \omega}{\omega_E}. \tag{8.17б}$$

Множитель при R в (8.17) учитывает зависимость коэффициента α_E от электрического поля (через $\omega_E \sim E^{2/3}$).

2. Рассмотрим выражение (8.17) для α_E в двух предельных случаях.

В первом случае $\beta > 0$ ($\omega < \omega_G$) и $\beta \gg 1$.

В этом случае в интеграле (8.17) существенны только большие положительные значения x, поэтому можно воспользоваться асимптотическим разложением[1])

$$Ai(x) \approx \frac{1}{2\sqrt{\pi}}x^{-1/4}\exp\left(-\frac{2}{3}x^{3/2}\right)\left(1 - \frac{5}{48}\frac{1}{x^{3/2}} + \ldots\right). \tag{8.18}$$

При $x \gg 1$ можно круглую скобку заменить единицей; подставляя

---

[1]) Джеффрис Г., Свирлс Б., с. 65, формула (20).

тогда (8.18) в интеграл (8.17), интегрируя по частям, получим

$$\int_{\beta}^{\infty} \frac{1}{4\pi} x^{-1/2} \exp\left(-\frac{4}{3}x^{3/2}\right) dx =$$

$$= -\left[\frac{1}{8\pi}\exp\left(-\frac{4}{3}x^{3/2}\right)\cdot\frac{1}{x}\Big|_{\beta}^{\infty} + \int_{\beta}^{\infty}\exp\left(-\frac{4}{3}x^{3/2}\right)\frac{1}{x^2}dx\right].$$

Интегрируя по частям интеграл в правой части, легко показать, что он в $\sim\beta^{3/2}$ раза меньше первого слагаемого; пренебрегая им, получим

$$\alpha_E = R\,(\hbar\omega_E)^{1/2}\,\frac{1}{8\beta}\exp\left(-\frac{4}{3}\beta^{3/2}\right). \tag{8.19}$$

Мы видим, что в этом случае в основном

$$\alpha_E \sim \exp\left[-\gamma\frac{(\omega_G-\omega)^{3/2}}{E}\right], \tag{8.19a}$$

где $\gamma = \dfrac{4\sqrt{2\mu\hbar}}{3e}$ — постоянная, не зависящая от частоты ω и электрического поля.

Во втором случае $\beta < 0$ $(\omega > \omega_G)$ и $|\beta| \gg 1$.

Асимптотические разложения функции Эйри при отрицательных, больших по абсолютной величине значениях $x = -\xi$ $(\xi \gg 1)$ имеют вид [1])

$$Ai(-\xi) \approx \frac{1}{\sqrt{\pi}}\xi^{-1/4}\left[\sin\left(\frac{2}{3}\xi^{3/2}+\frac{\pi}{4}\right) - \frac{5}{48}\xi^{-3/2}\cos\left(\frac{2}{3}\xi^{3/2}+\frac{\pi}{4}\right)\right]. \tag{8.20}$$

Так как при $\beta < 0$ в интеграле (8.17) x пробегает как положительные, так и отрицательные значения, включая нуль,— непосредственное применение формул (8.18) и (8.20) невозможно.

Выведем соотношение, которое может быть использовано в этом случае. Функция Эйри удовлетворяет дифференциальному уравнению [2])

$$\frac{d^2Ai(x)}{dx^2} = xAi(x). \tag{8.21}$$

Умножая обе части уравнения на $2dAi(x)/dx = 2Ai'$, получим

$$[Ai'^2]' = x[Ai^2]'. \tag{8.22}$$

Интегрируя обе части этого равенства от 0 до $\beta = -\xi$ и беря интеграл справа по частям, получим

$$-\int_0^{-\xi} Ai^2(x)\,dx = \xi Ai^2(-\xi) + [Ai'(-\xi)]^2. \tag{8.23}$$

---

[1]) Джеффрис Г., Свирлс Б., с. 65, формула (22).

[2]) Ландау Л. Д., Лифшиц Е. М. Квантовая механика, 1973, формула (24, 3).

Полагая $\beta = -\xi$, имеем для интеграла (8.17)

$$\int\limits_{-\xi}^{\infty} Ai^2(x)\,dx = \int\limits_{-\xi}^{0} Ai^2(x)\,dx + \int\limits_{0}^{\infty} Ai^2(x)\,dx = -\int\limits_{0}^{-\xi} Ai^2(x)\,dx + \\ + \int\limits_{0}^{\infty} Ai^2(x)\,dx = \xi Ai^2(-\xi) + [Ai'(-\xi)]^2 + \int\limits_{0}^{\infty} Ai^2(x)\,dx. \quad (8.24)$$

При $\xi \gg 1$ воспользуемся асимптотическим разложением (8.20); тогда в наинизшем приближении получим

$$\xi Ai^2(-\xi) + [Ai'(-\xi)]^2 = \frac{1}{\pi}\xi^{1/2}\sin^2\left(\frac{2}{3}\xi^{3/2} + \frac{\pi}{4}\right) + \\ + \frac{1}{\pi}\xi^{1/2}\cos^2\left(\frac{2}{3}\xi^{3/2} + \frac{\pi}{4}\right) = \frac{1}{\pi}\xi^{1/2} = \frac{1}{\pi}(-\beta)^{1/2} = \frac{1}{\pi}\left(\frac{\omega-\omega_G}{\omega_E}\right)^{1/2}. \quad (8.25)$$

Подставляя это в (8.24), получим для (8.17)

$$\alpha_E = R\left[(\hbar\omega - \varepsilon_G)^{1/2} + (\hbar\omega_E)^{1/2}\,\pi\int\limits_{0}^{\infty} Ai^2(x)\,dx\right]. \quad (8.26)$$

При электрическом поле $E \longrightarrow 0$ получим

$$\alpha_{E=0} = R\,(\hbar\omega - \varepsilon_G)^{1/2},$$

что совпадает с (2.26б).

3. Экспоненциальный вид коэффициента поглощения α_E (8.19а) можно понять и из наглядных соображений.

При приложении к полупроводнику однородного электрического поля, много меньшего атомных полей кристалла, энергетические зоны электрона наклоняются так, как это показано на рис. VII.5, где по вертикальной оси отложена энергия электрона ε, а по горизонтальной — пространственная координата x. Электрон с постоянной энергией может двигаться в валентной зоне между точками A и B, отражаясь от краев зоны (мы отвлекаемся от того, что валентная зона заполнена электронами, подчиняющимися принципу Паули). Однако из-за того, что электрон обладает волновыми свойствами, он может посредством туннельного эффекта преодолеть запрещенную зону BC и проникнуть в зону проводимости CD. Это явление называется *внутренней эмиссией в поле* или *эффектом Зинера*.

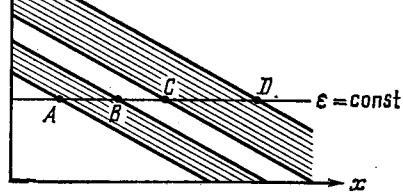


Рис. VII. 5.

Вероятность туннелирования экспоненциально падает с увеличением зоны BC, которая растет при увеличении ширины

запрещенной зоны ε_G и уменьшении электрического поля E. Расчет показывает, что вероятность туннелирования

$$W(B \to C) \sim \exp\left(-c\frac{\varepsilon_G^{3/2}}{E}\right), \tag{8.27}$$

где c — постоянная, зависящая от эффективных масс m_c и m_v. Выражение (8.27) подобно (8.19а) для α_E. Если $\omega > \omega_G$, то рождающиеся электрон и дырка при движении отражаются от барьера, созданного электрическим полем. Интерференция отраженной и падающей волн приводит к возникновению осцилляций коэффициента поглощения.

Учет кулоновского взаимодействия электрона и дырки меняет количественный характер поглощения света, однако качественные черты явления сохраняются (И. А. Меркулов, В. И. Перель, 1973). Так, если $\omega < \omega_G$, то экспоненциальный характер зависимости (8.19) сохраняется, однако изменяется коэффициент при $\beta^{3/2}$ в показателе экспоненты и появляется существенный дополнительный предэкспоненциальный множитель. Первое связано с изменением формы барьера при туннелировании, второе — с изменением вероятности найти электрон и дырку в данной точке из-за кулоновского взаимодействия.

Если кулоновское взаимодействие электрона и дырки приводит к связанному состоянию — экситону, то при частоте падающего света, близкой к порогу рождения экситона, в коэффициенте поглощения появляется линейчатый спектр. В электрическом поле эти линии поглощения испытывают штарк-эффект, аналогичный штарк-эффекту в атоме водорода; кроме того, у них появляется конечная ширина, равная обратной вероятности ионизации экситона в данном состоянии (И. А. Меркулов, 1974; А. Г. Аронов, А. С. Иоселевич, 1977).

Коэффициент поглощения света, с учетом кулоновского взаимодействия электрона и дырки, может быть вычислен и в случае $\omega > \omega_G$ (А. Г. Аронов, А. С. Иоселевич, 1977).

Если сравнить результаты п. 2 предыдущего параграфа с результатами теории Франка — Келдыша, то возникает вопрос: как согласовать эти результаты, когда магнитное поле стремится к нулю? Из (7.31) видно, что при $H \to 0$ сдвиг края поглощения, который равен $\frac{(m_c+m_v)c^2}{2}\left(\frac{E}{H}\right)^2 \to \infty$. Очевидно, что при $H \to 0$ нарушаются условия применимости формул, полученных в п. 2 предыдущего параграфа.

Можно показать, что теория, развитая в § 7, п. 2, справедлива при условии

$$\frac{(m_c+m_v)c^2}{2}\frac{E^2}{H^2} \ll \varepsilon_G \tag{8.28}$$

или

$$E/H \ll s/c, \tag{8.28a}$$

где скорость $s = \sqrt{\frac{2\varepsilon_G}{m_c + m_v}}$. Пока неравенство (8.28) выполнено, движение электрона и дырки носит финитный характер — имеет место квантование Ландау. При нарушении этого неравенства электрон и дырка движутся инфинитно, квантование Ландау отсутствует и имеет место эффект Франца — Келдыша в поперечном магнитном поле (А. Г. Аронов, Г. Е. Пикус, 1965).

ГЛАВА VIII

КИНЕТИЧЕСКОЕ УРАВНЕНИЕ И ВРЕМЯ РЕЛАКСАЦИИ ДЛЯ ЭЛЕКТРОНОВ ПРОВОДИМОСТИ В КРИСТАЛЛАХ

§ 1. Явления переноса и кинетическое уравнение Больцмана

1. В первых пяти параграфах гл. VI мы рассмотрели тепловые и магнитные свойства полупроводников и металлов, обусловленные свободными электронами и дырками, находящимися в состоянии статистического (термодинамического) равновесия. Как уже отмечалось выше, существенной особенностью систем, находящихся в статистическом равновесии, является то, что их свойства не зависят от механизма взаимодействия в системе.

Наряду с такими равновесными состояниями большой теоретический и практический интерес представляет изучение электронов (дырок) проводимости в неравновесном состоянии, когда они движутся в кристалле под действием приложенных внешних полей: электрического, магнитного, температурного. Такие процессы, связанные с перемещением электронов и дырок, называются *явлениями переноса* или *кинетическими эффектами*.

Если величины, описывающие явления переноса,— плотность электрического тока, тепловой поток, напряженность электрического поля и т. д.— не зависят от времени, то процесс называется *стационарным*. Для того чтобы при действии электрического поля, ускоряющего электроны, ток был стационарен, необходимо, чтобы электроны проводимости сталкивались (рассеивались) на каких-либо неоднородностях решетки (колебаниях атомов или дефектах кристалла) и отдавали бы накопленную в электрическом поле энергию. Как мы увидим дальше, в большинстве случаев столкновение (рассеяние) электрона практически можно рассматривать как упругое, и электрическое сопротивление будет определяться средней скоростью изменения составляющей импульса (скорости) электрона в направлении электрического поля при его рассеянии.

Важной особенностью неравновесных процессов является то, что они существенно зависят от механизма взаимодействия в

системе, в нашем случае — от взаимодействия электрона проводимости с колебаниями решетки или дефектами кристалла.

В гл. VI мы видели, что электроны в состоянии термодинамического равновесия в классическом случае описываются равновесной функцией распределения Больцмана $f_0(\mathscr{E}) = \exp\left(\frac{\zeta - \mathscr{E}}{k_0 T}\right)$, где полная энергия $\mathscr{E} = \frac{mv^2}{2} + \mathcal{U}(x, y, z)$ ($\mathcal{U}(x, y, z)$ — потенциальная энергия, которая у нас обычно равнялась нулю). Аналогично для электронов в неравновесном состоянии можно ввести неравновесную функцию распределения, имеющую тот смысл, что

$$f(v_x, v_y, v_z, x, y, z, t)\, dv_x\, dv_y\, dv_z\, dx\, dy\, dz = f(\boldsymbol{v}, \boldsymbol{r}, t)\, d^3v\, d^3r \quad (1.1)$$

равно числу электронов в момент t в точке $\boldsymbol{r}$ в объеме $d^3r = dx\, dy\, dz$, со скоростями, лежащими между v_x и $v_x + dv_x$ и т. д. [1]).

Такое описание электронов посредством одновременного задания их координат и скоростей (сопряженных импульсов) возможно только в той мере, в какой их движение подчиняется законам классической механики.

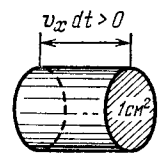


Рис. VIII.1.

Если известна функция $f(\boldsymbol{v}, \boldsymbol{r}, t)$, то можно вычислить плотность тока в точке $\boldsymbol{r}$ в момент t. На рис. VIII.1 изображена площадка в 1 $см^2$, перпендикулярная к плоскости рисунка и оси x, и цилиндр высотой $v_x\, dt$, построенный на этой площадке. Число $\boldsymbol{v}$-электронов внутри цилиндра равно $f(\boldsymbol{v}, \boldsymbol{r}, t)\, d^3v\, v_x\, dt$. Все эти электроны за время dt сместятся в направлении x на величину $v_x\, dt$ и, следовательно, пересекут площадку. Полное число электронов всех скоростей, пересекающих площадку за время dt, равно

$$dt \int_{-\infty}^{+\infty}\int\int f(\boldsymbol{v}, \boldsymbol{r}, t)\, v_x\, dv_x\, dv_y\, dv_z,$$

если учитывать электроны, пересекающие площадку как слева направо, так и справа налево. Так как каждый электрон имеет заряд $-e$, то плотность тока в направлении x равна

$$j_x = -e \int_{-\infty}^{+\infty}\int\int f(\boldsymbol{v}, \boldsymbol{r}, t)\, v_x\, dv_x\, dv_y\, dv_z. \quad (1.2)$$

Запишем неравновесную функцию распределения в виде

$$f(\boldsymbol{v}, \boldsymbol{r}, t) = f_0(\mathscr{E}) + f_1(\boldsymbol{v}, \boldsymbol{r}, t). \quad (1.3)$$

[1]) Такие электроны мы будем называть $\boldsymbol{v}$-электронами.

Так как $f_0(\mathscr{E})$—четная функция v_x (зависит от v_x^2), то интеграл (1.2) по dv_x от $f_0(\mathscr{E})\,v_x$ равен нулю и (1.2) дает

$$j_x = -e\int_{-\infty}^{+\infty}\!\!\iiint f_1(\boldsymbol{v}, \boldsymbol{r}, t)\,v_x\,dv_x\,dv_y\,dv_z. \tag{1.2a}$$

Основной задачей теории кинетических явлений является определение неравновесной функции распределения $f(\boldsymbol{v}, \boldsymbol{r}, t)$.

2. Выведем уравнение, которому удовлетворяет функция $f(\boldsymbol{v}, \boldsymbol{r}, t)$. Рассмотрим изменение числа $\boldsymbol{v}$-электронов за время dt в результате их движения в обычном (геометрическом) пространстве. На рис. VIII.2 представлен элемент объема $d^3\boldsymbol{r} = dx\,dy\,dz$. Рассмотрим изменение числа $\boldsymbol{v}$-электронов внутри этого объема за счет прихода электронов сквозь левую грань $dy\,dz$ и ухода электронов сквозь правую грань $dy\,dz$ (мы предполагаем, что $v_x > 0$). Число $\boldsymbol{v}$-электронов, входящих за время dt через левую грань, равно $f(\boldsymbol{v}, x, y, z, t) \times d^3v\,dy\,dzv_x\,dt$, число $\boldsymbol{v}$-электронов, уходящих за время dt через правую грань, равно $f(\boldsymbol{v}, x+dx, y, z, t)\,d^3v\,dy\,dzv_x\,dt$. Увеличение числа $\boldsymbol{v}$-электронов в объеме d^3r за счет этого процесса равно

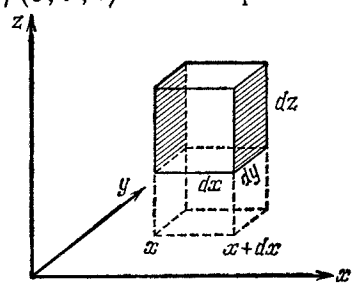


Рис. VIII. 2.

$$f(\boldsymbol{v}, x, y, z, t)\,d^3v\,dy\,dzv_x\,dt - f(\boldsymbol{v}, x+dx, y, z, t)\,d^3v\,dy\,dzv_x\,dt = \\ = -v_x\{f(\boldsymbol{v}, x+dx, y, z, t) - f(\boldsymbol{v}, x, y, z, t)\}\,dy\,dz\,d^3v\,dt = \\ = -v_x\frac{\partial f}{\partial x}\,dx\,dy\,dz\,d^3v\,dt.$$

Увеличение числа $\boldsymbol{v}$-электронов в объеме $d^3\boldsymbol{r}$ за счет движения электронов сквозь все 6 граней объема $d^3\boldsymbol{r}$, равно

$$-\left(v_x\frac{\partial f}{\partial x} + v_y\frac{\partial f}{\partial y} + v_z\frac{\partial f}{\partial z}\right)d^3\boldsymbol{v}\,d^3\boldsymbol{r}\,dt = -(\boldsymbol{v}\nabla_r f)\,d^3v\,d^3r\,dt = \\ = -\boldsymbol{v}\frac{\partial f}{\partial \boldsymbol{r}}\,d^3\boldsymbol{v}\,d^3\boldsymbol{r}\,dt. \tag{1.4}$$

Аналогично можно рассмотреть изменение числа $\boldsymbol{v}$-электронов в объеме $d^3v = dv_x\,dv_y\,dv_z$ в результате их «движения» в $\boldsymbol{v}$-пространстве со «скоростями» $\dot{v}_x$, $\dot{v}_y$ и $\dot{v}_z$. Увеличение числа $\boldsymbol{v}$-электронов при движении в $\boldsymbol{v}$-пространстве равно

$$-(\dot{\boldsymbol{v}}\nabla_v f)\,d^3\boldsymbol{v}\,d^3\boldsymbol{r}\,dt = -\dot{\boldsymbol{v}}\frac{\partial f}{\partial \boldsymbol{v}}\,d^3\boldsymbol{v}\,d^3\boldsymbol{r}\,dt = -\frac{1}{m}(\boldsymbol{F}\nabla_v f)\,d^3\boldsymbol{v}\,d^3\boldsymbol{r}\,dt, \tag{1.4a}$$

так как ускорение $\dot{\boldsymbol{v}} = \frac{d\boldsymbol{v}}{dt} = \frac{1}{m}\boldsymbol{F}(\boldsymbol{r}, t)$, где $\boldsymbol{F}(\boldsymbol{r}, t)$—сила, действующая на электрон в точке $\boldsymbol{r}$ в момент t.

Кроме того, число $\boldsymbol{v}$-электронов изменяется в результате столкновений электронов с колебаниями решетки (фононами) и дефектами кристалла. Каждое столкновение $\boldsymbol{v}$-электрона уводит его из объема $d^3\boldsymbol{v}$, так как при столкновении резко меняется его скорость. В дальнейшем мы будем рассматривать упругие столкновения, при которых меняется только направление скорости, но не ее абсолютная величина; так что если после столкновения скорость равна $\boldsymbol{v}'$, то $v'=v$.

Пусть $W(\boldsymbol{v}, \boldsymbol{v}')\, d^3v'\, dt$ — вероятность электрону со скоростью $\boldsymbol{v}$ за время dt упруго рассеяться и превратиться в $\boldsymbol{v}'$-электрон. В принципе вероятность $W(\boldsymbol{v}, \boldsymbol{v}')$ может зависеть от $\boldsymbol{r}$ и t.

Полное число $\boldsymbol{v}$-электронов, исчезающих за время dt в результате столкновений, равно

$$-\int_{\boldsymbol{v}'} [f(\boldsymbol{v}, \boldsymbol{r}, t)\, d^3v\, d^3r W(\boldsymbol{v}, \boldsymbol{v}')\, dt]\, d^3v',$$

где интегрирование ведется по всем значениям скорости $\boldsymbol{v}'$.

С другой стороны, число $\boldsymbol{v}$-электронов будет возрастать ввиду превращения всевозможных $\boldsymbol{v}'$-электронов в $\boldsymbol{v}$-электроны в результате столкновений (в том же объеме d^3r). Это возрастание числа $\boldsymbol{v}$-электронов за время dt равно

$$\int_{\boldsymbol{v}'} [f(\boldsymbol{v}', \boldsymbol{r}, t)\, d^3r W(\boldsymbol{v}', \boldsymbol{v})\, d^3v\, dt]\, d^3v',$$

где интегрирование ведется по всем значениям скорости $\boldsymbol{v}'$.

В итоге возрастание числа $\boldsymbol{v}$-электронов за время dt из-за столкновений равно

$$d^3v\, d^3r\, dt \int_{\boldsymbol{v}'} [f(\boldsymbol{v}', \boldsymbol{r}, t) W(\boldsymbol{v}', \boldsymbol{v}) - f(\boldsymbol{v}, \boldsymbol{r}, t) W(\boldsymbol{v}, \boldsymbol{v}')]\, d^3v'. \qquad (1.4б)$$

С другой стороны, возрастание числа $\boldsymbol{v}$-электронов за время dt равно

$$f(\boldsymbol{v}, \boldsymbol{r}, t+dt)\, d^3v\, d^3r - f(\boldsymbol{v}, \boldsymbol{r}, t)\, d^3v\, d^3r = \frac{\partial f}{\partial t}\, d^3v\, d^3r\, dt. \qquad (1.5)$$

Приравнивая (1.5) сумме (1.4), (1.4а) и (1.4б), получим

$$\frac{\partial f}{\partial t} = -(\boldsymbol{v}\nabla_r f) - \frac{1}{m}(\boldsymbol{F}\nabla_v f) + \\ + \int \{f(\boldsymbol{v}', \boldsymbol{r}, t) W(\boldsymbol{v}', \boldsymbol{v}) - f(\boldsymbol{v}, \boldsymbol{r}, t) W(\boldsymbol{v}, \boldsymbol{v}')\}\, d^3v'. \qquad (1.6)$$

Это уравнение называется *кинетическим уравнением Больцмана*.

Так называемый *полевой* член

$$\left(\frac{\partial f}{\partial t}\right)_{\text{пол}} = -(\boldsymbol{v}\nabla_r f) - \frac{1}{m}(\boldsymbol{F}\nabla_v f) \qquad (1.7)$$

определяет скорость изменения функции распределения f в результате непрерывного движения электронов в $\boldsymbol{r}$ и $\boldsymbol{v}$-простран-

стве, а член *столкновений*

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = \int \{f(\boldsymbol{v}', \boldsymbol{r}, t)\, W(\boldsymbol{v}', \boldsymbol{v}) - f(\boldsymbol{v}, \boldsymbol{r}, t)\, W(\boldsymbol{v}, \boldsymbol{v}')\}\, d^3v' \qquad (1.7a)$$

определяет скорость изменения f в результате столкновения (рассеяния) электронов.

В стационарном случае

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{\text{пол}} + \left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = 0 \qquad (1.8)$$

или

$$\boldsymbol{v}\nabla_r f + \frac{1}{m}\boldsymbol{F}\nabla_v f = \int \{f(\boldsymbol{v}', \boldsymbol{r})\, W(\boldsymbol{v}', \boldsymbol{v}) - f(\boldsymbol{v}, \boldsymbol{r})\, W(\boldsymbol{v}, \boldsymbol{v}')\}\, d^3v'. \qquad (1.8a)$$

Так как в правой части (1.8а) стоит интеграл от неизвестной функции $f(\boldsymbol{v}')$, то кинетическое уравнение Больцмана — интегродифференциальное уравнение. Конечно, для его решения необходимо знать силу $\boldsymbol{F}$ и вероятность перехода $W(\boldsymbol{v}, \boldsymbol{v}')$.

3. В равновесном случае функция распределения $f = f_0(\mathscr{E}) = \exp\left\{\frac{\zeta - \mathscr{E}}{k_0 T}\right\}$, где полная энергия $\mathscr{E} = \varepsilon + \mathcal{U}$, т. е. равна сумме кинетической энергии $\varepsilon = mv^2/2$ и потенциальной энергии $\mathcal{U}(\boldsymbol{r})$. В этом случае левая часть уравнения (1.8а) равна нулю. В самом деле,

$$\boldsymbol{v}\nabla_r f_0(\varepsilon + \mathcal{U}) = \frac{1}{k_0 T} f_0 \boldsymbol{v}(-\nabla_r \mathcal{U}) = \frac{1}{k_0 T} f_0(\boldsymbol{vF})$$

и

$$\frac{1}{m}\boldsymbol{F}\nabla_v f_0(\varepsilon + \mathcal{U}) = -\frac{1}{m}\boldsymbol{F}\frac{m\boldsymbol{v}}{k_0 T} f_0 = -\frac{1}{k_0 T} f_0(\boldsymbol{vF}).$$

Таким образом,

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = \int \{f_0(\mathscr{E}')\, W(\boldsymbol{v}', \boldsymbol{v}) - f_0(\mathscr{E})\, W(\boldsymbol{v}, \boldsymbol{v}')\}\, d^3v' =$$

$$= f_0(\mathscr{E}) \int \{W(\boldsymbol{v}', \boldsymbol{v}) - W(\boldsymbol{v}, \boldsymbol{v}')\}\, d^3v' = 0, \qquad (1.9)$$

так как $\mathscr{E} = \mathscr{E}'$. При произвольном $\boldsymbol{v}$ интеграл может равняться нулю только в случае

$$W(\boldsymbol{v}, \boldsymbol{v}') = W(\boldsymbol{v}', \boldsymbol{v}). \qquad (1.10)$$

Последнее равенство есть следствие общего принципа детального равновесия, согласно которому вероятности прямого и обратного процессов одинаковы. При вычислении вероятностей перехода в квантовой механике равенство (1.10) является прямым следствием ее законов.

Если на электроны действует только электрическое поле $\boldsymbol{E}$, направленное вдоль оси x, то, как будет показано ниже, поправка

к функции распределения может быть представлена в виде [1])

$$f_1 = -\frac{\partial f_0}{\partial \varepsilon}\chi(\varepsilon)\, v_x, \tag{1.11}$$

где $\chi(\varepsilon)$ — некоторая функция энергии $\varepsilon = \frac{m}{2}(v_x^2 + v_y^2 + v_z^2)$. Подставляя (1.11) в (1.7a), считая столкновения упругими ($\varepsilon' = \varepsilon$) и учитывая равенство (1.10), получим

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -f_1 \int W(\boldsymbol{v}, \boldsymbol{v}')\left(1 - \frac{v_x'}{v_x}\right) d^3v'. \tag{1.12}$$

Так как мы предполагаем, что рассеяние электронов упругое, то

$$W(\boldsymbol{v}, \boldsymbol{v}') = W_0(v, \theta)\,\delta(v - v'), \tag{1.12а}$$

где $\delta(v - v')$ учитывает, что при рассеянии $v = v'$ и θ — угол между направлениями скоростей $\boldsymbol{v}$ и $\boldsymbol{v}'$.

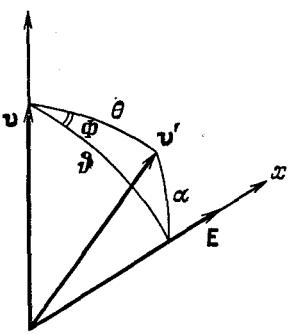


Рис. VIII. 3.

Выбирая $\boldsymbol{v}$ в качестве полярной оси в пространстве скоростей $\boldsymbol{v}'$, как это показано на рис. VIII.3, имеем

$$d^3v' = v'^2 \sin\theta\, d\theta\, d\Phi = v'^2\, d\Omega, \tag{1.12б}$$

где θ и Φ — полярный и азимутальный углы, определяющие направление вектора $\boldsymbol{v}'$, а $d\Omega = \sin\theta\, d\theta\, d\Phi$ — телесный угол в направлении $\boldsymbol{v}'$. Ось x совпадает с направлением электрического поля $\boldsymbol{E}$. Имеем

$$v_x' = v'\cos\alpha = v\cos\alpha, \quad v_x = v\cos\vartheta. \tag{1.12в}$$

Из сферической тригонометрии известно [2]), что

$$\cos\alpha = \cos\vartheta\cos\theta + \sin\vartheta\sin\theta\cos\Phi. \tag{1.12г}$$

Подставляя (1.12а)—(1.12г) в (1.12), получим

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -f_1 \int W(\theta)(1 - \cos\theta)\, d\Omega, \tag{1.13}$$

где $W(\theta)\, d\Omega = v^2 W_0(v, \theta)\, d\Omega$ — вероятность упругого рассеяния за 1 *сек* электрона со скоростью v в телесный угол $d\Omega$. Так как интеграл в (1.12) имеет размерность, обратную времени, то можно

---

[1]) Т. е. в виде произведения некоторой функции от энергии ε на v_x; множитель $\left(-\frac{\partial f_0}{\partial \varepsilon}\right)$ выделяется для удобства дальнейших расчетов.

[2]) Кочин Н. Е., гл. I.

ввести *время релаксации* τ, положив

$$\frac{1}{\tau} = \int W(\theta)(1-\cos\theta)\, d\Omega. \tag{1.13a}$$

Время релаксации зависит только от скорости электрона v (или его энергии ε) и механизма его рассеяния. Из (1.13) и (1.13а) следует, что

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -\frac{f_1}{\tau} = -\frac{f-f_0}{\tau}. \tag{1.14}$$

Наглядный смысл времени релаксации можно выяснить, рассматривая установление статистического равновесия в однородной системе, в которой в начальный момент времени $t=0$ распределение по скоростям не было равновесным и на которую не действуют силы. Полагая $\nabla_r f = 0$ и $\boldsymbol{F}=0$, имеем на основании выражений (1.6), (1.7а), (1.14)

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = \frac{\partial f}{\partial t} = -\frac{f-f_0}{\tau}.$$

Интегрируя последнее равенство, получим

$$(f-f_0) = (f-f_0)_{t=0}\, e^{-t/\tau}. \tag{1.14a}$$

Мы видим, что τ есть время, в течение которого разность $(f-f_0)$ при выключении внешнего поля уменьшается в e раз. Так как стремление к равновесию в системе происходит в результате столкновений электронов (с колебаниями решетки или дефектами кристалла) и при этом достаточно нескольких столкновений для того, чтобы электроны пришли к равновесному состоянию, то время релаксации τ порядка времени свободного пробега электрона. Мы определим среднюю длину свободного пробега электрона

$$l = v\tau. \tag{1.15}$$

Как мы увидим ниже, во многих практически интересных случаях поправка к равновесной функции распределения $f_1 \ll f_0$, поэтому с точностью до величин первого порядка f_1 можно в левую часть уравнения (1.8а) подставить вместо f равновесную функцию f_0. При наличии только электрического поля из (1.8а) и (1.14)

$$f_1 = \frac{e}{m} E \frac{\partial f_0}{\partial v_x}\tau = -eE\frac{f_0}{k_0 T}\tau v_x. \tag{1.16}$$

Таким образом, если нам известно время релаксации τ, то по формуле (1.2а) можно определить плотность тока. Одновременно мы видим, что f_1 действительно имеет вид (1.11), где $\chi(\varepsilon) = -eE\tau(v)$.

4. Вычислим время релаксации по формуле (1.13а) для электронов (дырок), сталкивающихся (рассеивающихся) с примесными

ионами, распределенными с постоянной средней плотностью по объему проводника. Рассмотрим по законам классической механики движение дырки в кулоновском поле донора:

$$\mathcal{U}(r) = \frac{e^2}{\varepsilon_0 r}. \tag{1.17}$$

Здесь e — элементарный заряд, ε_0 — диэлектрическая постоянная.

Можно показать[1]), что дырка движется по гиперболе и что ее «прицельное расстояние»

$$b = \frac{e^2}{\varepsilon_0 m v^2} \operatorname{ctg} \frac{\theta}{2}, \tag{1.18}$$

где θ — угол рассеяния дырки (рис. VIII.4). Выражение (1.18) справедливо и в случае рассеяния электрона на доноре, когда потенциальная энергия притяжения равна (1.17) со знаком минус. Конечно, в этом случае движение по гиперболе имеет место только в том случае, если полная энергия электрона $\frac{mv^2}{2} + \mathcal{U}(r) > 0$[2]).

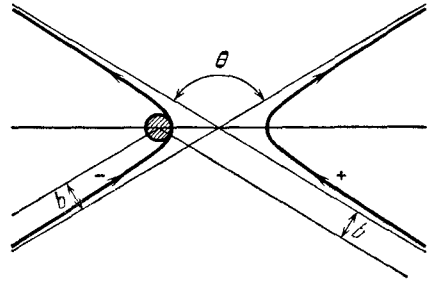


Рис. VIII. 4.

Дифференциальное эффективное сечение $\sigma(\theta)$ по определению равно

$$\sigma(\theta)\, d\Omega = \frac{\text{число электронов, отклоненных на угол } \theta \text{ в телесный угол } d\Omega}{\text{число электронов, упавших на 1 } см^2 \text{ за то же время}} \tag{1.19}$$

и имеет размерность площади ($см^2$).

Все частицы, движущиеся параллельно оси x и падающие на кольцо площадью $2\pi b\, db$, отклонятся на угол θ в пределах телесного угла $d\Omega = 2\pi \sin\theta\, d\theta$, где связь между $|db|$ и $d\theta$ определяется из (1.18).

Из (1.19) следует, что

$$\sigma(\theta)\, 2\pi \sin\theta\, d\theta = 2\pi b\, |db|. \tag{1.19а}$$

Из (1.18) имеем

$$db = -\frac{e^2}{2\varepsilon_0 m v^2} \frac{1}{\sin^2 \frac{\theta}{2}} d\theta. \tag{1.18а}$$

Из выражений (1.19а), (1.18) и (1.18а) получим формулу

[1]) Борн М., Атомная физика. М.: Мир, 1965, с. 384.

[2]) На рис. VIII.4 правая ветвь гиперболы изображает траекторию дырки, а левая ветвь — траекторию электрона.

Резерфорда

$$\sigma(\theta) = \left(\frac{e^2}{2\varepsilon_0 m v^2}\right)^2 \frac{1}{\sin^4 \frac{\theta}{2}}, \tag{1.20}$$

использованную им при изучении рассеяния α-частиц на ядрах тяжелых элементов.

Определим интегральное сечение как

$$\sigma = \int \sigma(\theta)\, d\Omega = 2\pi \int_0^{\pi} \sigma(\theta) \sin\theta\, d\theta, \tag{1.21}$$

т. е. как полное число рассеянных частиц в расчете на единичную плотность потока падающих частиц.

Легко показать, что для кулоновского поля, когда дифференциальное сечение имеет вид (1.20), интегральное сечение σ равно бесконечности. Это является специфической особенностью кулоновского потенциала (1.17) и обусловливается медленностью его спадания. Установим связь между $\sigma(\theta)$ и вероятностью $W(\boldsymbol{v}, \boldsymbol{v}')$, которая при упругом рассеянии зависит от угла θ между скоростями $\boldsymbol{v}'$ и $\boldsymbol{v}$ и их абсолютным значением.

Пусть в объеме V имеется кулоновский центр и N электронов, движущихся по всем направлениям со скоростями v. В 1 *сек* на этот центр упадет поток электронов, равный $(N/V)\, v$ (мы для электронов каждого направления выберем площадку в 1 *см*^{2}, перпендикулярную к их движению). Полное число электронов, рассеянных в 1 *сек* на угол между θ и $\theta + d\theta$, равно $(N/V)\, v\sigma(\theta)\, d\Omega$.

С другой стороны, это же число равно $NW(\theta)\, d\Omega$; таким образом,

$$W(\theta) = v\sigma(\theta)/V. \tag{1.22}$$

Легко проверить, что правая часть равенства имеет размерность *сек*$^{-1}$.

Подставляя (1.22) в (1.13а) и заменяя $d\Omega' = 2\pi \sin\theta d\theta$, получим

$$\frac{1}{\tau} = \frac{2\pi v}{V} \int \sigma(\theta)\,(1 - \cos\theta) \sin\theta\, d\theta. \tag{1.23}$$

Если в объеме имеется N независимо рассеивающих ионов, то

$$\frac{1}{\tau} = 2\pi v n_I \int \sigma(\theta)\,(1 - \cos\theta) \sin\theta\, d\theta, \tag{1.23а}$$

где $n_I = N_I/V$ — концентрация ионов.

Если подставить сюда вместо $\sigma(\theta)$ выражение (1.20), то интеграл на нижнем пределе при $\theta = 0$ логарифмически расходится, что обусловлено (так же как и для расходимости интегрального сечения σ) медленностью убывания кулоновского потенциала.

Расходимость выражения (1.23а) может быть устранена, если мы в том или ином виде учтем поле, создаваемое остальными

электронами проводимости, которые экранируют поля ионов и «обрезают» действие кулоновского потенциала. С этой точки зрения представляется естественным сферу действия каждого рассеивающего центра ограничить половиной среднего расстояния между ионами. Тогда наибольшее значение прицельного расстояния $b_{\max}=\frac{1}{2}n_I^{-1/3}$; наименьшее значение угла рассеяния $\theta_{\min}$ определяется из формулы (1.18):

$$\operatorname{ctg}\frac{\theta_{\min}}{2}=\frac{\varepsilon_0 m v^2}{2n_I^{1/3}e^2}. \tag{1.24}$$

Из выражений (1.23а) и (1.20) следует

$$\frac{1}{\tau}=2\pi v n_I\left(\frac{e^2}{2\varepsilon_0 m v^2}\right)^2\int\limits_{\theta_{\min}}^{\pi}\frac{(1-\cos\theta)\sin\theta\,d\theta}{\sin^4(\theta/2)}.$$

Заменяя $1-\cos\theta$ и $\sin\theta$ на $2\sin^2(\theta/2)$ и $2\sin(\theta/2)\cos\theta/2$ и используя равенство (1.24), получим

$$\frac{1}{\tau}=\frac{2\pi n_I e^4}{\varepsilon_0^2 m^2 v^3}\lg\left\{1+\left[\frac{\varepsilon_0 m v^2}{2e^2 n_I^{1/3}}\right]^2\right\}. \tag{1.25}$$

Это выражение для времени релаксации τ при рассеянии носителей тока на ионах примеси часто называют *формулой Конвелл — Вейскопфа* (1946).

Так как логарифм в (1.25) — медленно меняющаяся функция v, то практически время релаксации

$$\tau\sim v^3\sim\varepsilon^{3/2}. \tag{1.25а}$$

Используя выражения (1.25) и (1.16), можно по формуле (1.2а) вычислить плотность тока, а следовательно, и удельную электропроводность.

§ 2. Кинетическое уравнение для электронов в кристалле

1. В предыдущем параграфе предполагалось, что движение электронов подчиняется законам классической механики, поэтому их состояние описывалось в $\boldsymbol{r}$- и $\boldsymbol{v}$-пространствах и считалось, что

$$\frac{d\boldsymbol{v}}{dt}=\frac{1}{m}\boldsymbol{F}. \tag{2.1}$$

Последнее уравнение справедливо и в квазиклассическом приближении, если энергия электрона $\varepsilon=h^2k^2/2m^*=m^*v^2/2$, где m^* — эффективная масса, а $\boldsymbol{v}=\hbar\boldsymbol{k}/m^*$.

Таким образом, если энергия электрона $\varepsilon(\boldsymbol{k})$ имеет приведенный выше вид, то применимы все формулы предыдущего параграфа при замене m на m^*.

В общем случае произвольного закона дисперсии $\varepsilon(\boldsymbol{k})$ скорость электрона в кристалле (IV.3.32) равна

$$\boldsymbol{v}=\frac{1}{\hbar}\nabla_{\boldsymbol{k}}\varepsilon(\boldsymbol{k})=\frac{1}{\hbar}\frac{\partial\varepsilon(\boldsymbol{k})}{\partial k} \tag{2.2}$$

и не равна $\hbar\boldsymbol{k}/m^*$; поэтому возникает вопрос: какой вид имеет кинетическое уравнение (1.8а) в этом случае?

В квазиклассическом приближении электрону может быть приписана траектория, по которой он движется со скоростью (2.2). Волновой вектор электрона $\boldsymbol{k}$, характеризующий его квантовое состояние, удовлетворяет в квазиклассическом приближении уравнению (IV.3.35):

$$\frac{d\boldsymbol{k}}{dt}=\frac{1}{\hbar}\boldsymbol{F}. \tag{2.3}$$

Рассматривая электроны в $\boldsymbol{r}$- и $\boldsymbol{k}$-пространстве, введем функцию распределения $f(\boldsymbol{k}, \boldsymbol{r}, t)$, так что

$$f(\boldsymbol{k}, \boldsymbol{r}, t)\frac{d^3k}{4\pi^3} \tag{2.4}$$

равно числу электронов в момент t в точке $\boldsymbol{r}$ в единице объема с составляющими волнового вектора от k_x до k_x+dk_x и т. д. ($d^3k=dk_x\,dk_y\,dk_z$).

В состоянии статистического равновесия (VI.2.1)

$$f=f_0(\varepsilon)=\left(e^{\frac{\varepsilon-\zeta}{k_0T}}+1\right)^{-1}, \tag{2.5}$$

где ε—заданная функция $\boldsymbol{k}$.

Так как вместо уравнения (2.1) мы имеем теперь (2.3), то кинетическое уравнение (1.8а), если на электрон действует электрическое и магнитное поле $\boldsymbol{E}$ и $\boldsymbol{H}$, будет иметь вид

$$\boldsymbol{v}\nabla_{\boldsymbol{r}}f-\frac{e}{\hbar}\left(\boldsymbol{E}+\frac{1}{c}[\boldsymbol{v}\boldsymbol{H}]\right)\nabla_{\boldsymbol{k}}f=\left(\frac{\partial f}{\partial t}\right)_{\text{ст}}, \tag{2.6}$$

где $\boldsymbol{v}$ определяется выражением (2.2). Член столкновений $(\partial f/\partial t)_{\text{ст}}$ может быть выражен аналогично тому, как это было сделано в предыдущем параграфе. Пусть $W(\boldsymbol{k}, \boldsymbol{k}')$—вероятность электрону за 1 *сек* перейти из состояния $\boldsymbol{k}$ в состояние $\boldsymbol{k}'$; тогда аналогично (1.7а)

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}}=\sum_{\boldsymbol{k}'}\{W(\boldsymbol{k}', \boldsymbol{k})f(\boldsymbol{k}')[1-f(\boldsymbol{k})]-W(\boldsymbol{k}, \boldsymbol{k}')f(\boldsymbol{k})[1-f(\boldsymbol{k}')]\}, \tag{2.7}$$

где учтен принцип Паули, т. е. вероятность перехода $\boldsymbol{k}\to\boldsymbol{k}'$ положена пропорциональной $[1-f(\boldsymbol{k}')]$ (вероятности того, что состояние $\boldsymbol{k}'$ свободно).

Из принципа детального равновесия следует, что в статистическом равновесии потоки электронов $\boldsymbol{k} \to \boldsymbol{k}'$ и обратный должны быть равны, т. е.

$$W(\boldsymbol{k}', \boldsymbol{k}) f_0(\varepsilon')[1-f_0(\varepsilon)] = W(\boldsymbol{k}, \boldsymbol{k}') f_0(\varepsilon)[1-f_0(\varepsilon')].$$

Используя явный вид функции распределения Ферми, получим

$$W(\boldsymbol{k}', \boldsymbol{k}) e^{\varepsilon/k_0 T} = W(\boldsymbol{k}, \boldsymbol{k}') e^{\varepsilon'/k_0 T}. \tag{2.7а}$$

При упругом рассеянии $\varepsilon = \varepsilon'$ и

$$W(\boldsymbol{k}', \boldsymbol{k}) = W(\boldsymbol{k}, \boldsymbol{k}'). \tag{2.7б}$$

В этом случае

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = \sum_{\boldsymbol{k}'} \{W(\boldsymbol{k}', \boldsymbol{k}) f(\boldsymbol{k}') - W(\boldsymbol{k}, \boldsymbol{k}') f(\boldsymbol{k})\} =$$

$$= \sum_{\boldsymbol{k}'} W(\boldsymbol{k}, \boldsymbol{k}') \{f(\boldsymbol{k}') - f(\boldsymbol{k})\}, \tag{2.8}$$

т. е. член столкновений получается таким же, как и без учета принципа Паули. Это обстоятельство связано с тем, что при учете принципа Паули увеличивается на одну и ту же величину как число электронов, переходящих из состояния $\boldsymbol{k}$ в $\boldsymbol{k}'$, так и обратно из $\boldsymbol{k}'$ в $\boldsymbol{k}$.

Как мы увидим в § 2 следующей главы, неравновесная функция распределения может быть представлена в виде

$$f(\boldsymbol{k}) = f_0(\varepsilon) + f_1(\boldsymbol{k}) = f_0(\varepsilon) - \frac{\partial f_0}{\partial \varepsilon} \chi(\varepsilon) \boldsymbol{k}, \tag{2.9}$$

где $\chi(\varepsilon)$ — неизвестная векторная функция, зависящая от энергии электрона ε [1]).

Будет показано, что неравновесная добавка имеет указанный вид:

$$f_1(\boldsymbol{k}) = -\frac{\partial f_0}{\partial \varepsilon} \chi(\varepsilon) \boldsymbol{k}, \tag{2.9а}$$

как при наличии электрического и магнитного поля, так и при наличии градиента температуры, если энергия $\varepsilon \sim k^2$ и τ зависит только от ε.

Так как при упругом рассеянии $\varepsilon' = \varepsilon$, то

$$f(\boldsymbol{k}') - f(\boldsymbol{k}) = f_0(\varepsilon') - \frac{\partial f_0}{\partial \varepsilon'} \chi(\varepsilon') \boldsymbol{k}' - f_0(\varepsilon) + \frac{\partial f_0}{\partial \varepsilon} \chi(\varepsilon) \boldsymbol{k} =$$

$$= \frac{\partial f_0}{\partial \varepsilon} \chi(\varepsilon) \boldsymbol{k} \left[1 - \frac{k'_\chi}{k_\chi}\right] = f_1(\boldsymbol{k}) \frac{\Delta k_\chi}{k_\chi}, \tag{2.10}$$

---

[1]) Множитель $\partial f_0/\partial \varepsilon$, тоже зависящий только от энергии ε, введен для удобства выкладок.

где k_χ—проекция $\boldsymbol{k}$ на вектор $\boldsymbol{\chi}$ и $\Delta k_\chi = k'_\chi - k_\chi$. Из (2.8) и (2.10)

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -f_1(\boldsymbol{k}) \sum_{\boldsymbol{k}'} W(\boldsymbol{k}, \boldsymbol{k}') \left[1 - \frac{k'_\chi}{k_\chi}\right]. \tag{2.11}$$

Введем время релаксации τ, полагая

$$\frac{1}{\tau(\boldsymbol{k})} = \sum_{\boldsymbol{k}'} W(\boldsymbol{k}, \boldsymbol{k}') \left[1 - \frac{k'_\chi}{k_\chi}\right] = -\sum_{\boldsymbol{k}'} W(\boldsymbol{k}, \boldsymbol{k}') \frac{\Delta k_\chi}{k_\chi} \tag{2.12}$$

тогда (2.11) можно записать в виде

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -\frac{f_1(\boldsymbol{k})}{\tau} = -\frac{f - f_0}{\tau}. \tag{2.11а}$$

Если поверхности постоянной энергии электрона $\varepsilon(\boldsymbol{k}) = \text{const}$ не сферы, а эллипсоиды, то, строго говоря, необходимо учитывать анизотропию в рассеянии электрона, и в этом случае время релаксации τ, если его можно ввести,— не скаляр, а тензор 2-го ранга [1]).

В случае сложного закона дисперсии энергии $\varepsilon(\boldsymbol{k})$, например такого, как для дырок в p-Ge и p-Si (IV.15, а, б) вообще невозможно строго ввести время релаксации; поэтому в этом случае теория кинетических явлений становится очень сложной.

2. В следующей главе будут рассмотрены с помощью кинетического уравнения (2.6) различные явления переноса в общем случае при наличии электрического и магнитного полей и градиента температуры. Здесь мы в иллюстративных целях вычислим электропроводность при наличии одного электрического поля $\boldsymbol{E}$.

При наличии одного электрического поля $\boldsymbol{E}$ получим из (2.6) и (2.11а), положив с точностью до первого порядка по $f_1 \sim E$, $f = f_0$,

$$-\frac{e}{\hbar} \boldsymbol{E} \nabla_k f_0 = -\frac{f_1(\boldsymbol{k})}{\tau(\boldsymbol{k})}. \tag{2.13}$$

Используя (2.2), получим из (2.13)

$$f_1(\boldsymbol{k}) = e\tau(\boldsymbol{k}) \frac{\partial f_0}{\partial \varepsilon} (\boldsymbol{v}\boldsymbol{E}). \tag{2.13а}$$

Плотность электрического тока аналогично (1.2) равна

$$\boldsymbol{j} = -\frac{e}{4\pi^3} \int \boldsymbol{v} f(\boldsymbol{k}) d^3k = -\frac{e}{4\pi^3} \int \boldsymbol{v} f_1(\boldsymbol{k}) d^3k =$$

$$= -\frac{e^2}{4\pi^3} \int \tau(\boldsymbol{k}) \frac{\partial f_0}{\partial \varepsilon} \boldsymbol{v} (\boldsymbol{v}\boldsymbol{E}) d^3k, \tag{2.14}$$

---

[1]) Херринг К., Фогт Э. Проблемы физики полупроводников, М.: ИЛ, 1957, с. 567; Самойлович А. Г., Коренблит И. Я., Даховский И. В., Искра В. Д.— Физика твердого тела, 1961, № 3, с. 3285.

так как ток, соответствующий равновесной функции f_0, равен нулю. Обозначая индексами i и l составляющие $\boldsymbol{v}$ и $\boldsymbol{E}$ в прямоугольной системе координат, имеем

$$j_l = -\frac{e^2}{4\pi^3}\int \tau(\boldsymbol{k})\frac{\partial f_0}{\partial \varepsilon} v_l \sum_i v_i E_i\, d^3k = \sum_i \sigma_{li} E_i. \qquad (2.14а)$$

Тензор электропроводности

$$\sigma_{li} = -\frac{e^2}{4\pi^3}\int \tau(\boldsymbol{k})\frac{\partial f_0}{\partial \varepsilon} v_l v_i\, d^3k, \qquad (2.14б)$$

где v_l, как функция $\boldsymbol{k}$, определяется через (2.2).

Легко показать, что если время релаксации τ зависит только от абсолютной величины $|\boldsymbol{k}|$, то тензор электропроводности (2.14б) вырождается в скаляр.

§ 3. Рассеяние электронов на акустических колебаниях решетки

1. Мы видели, что стационарному состоянию электрона в периодическом поле кристалла соответствует не зависящая от времени скорость (2.2). Так как электрон обладает зарядом $-e$, то постоянной скорости соответствует незатухающий ток в отсутствие электрического поля. Можно сказать, что сопротивление идеального кристалла равно нулю. Конечное сопротивление кристалла, с которым связано конечное время релаксации τ или длина свободного пробега $l = v\tau$, обусловливается любыми отступлениями поля кристалла от строгой периодичности. Одной из существенных причин нарушения периодичности являются тепловые колебания атомов решетки, рассмотренные в гл. III.

В этом параграфе мы вычислим время релаксации, связанное с рассеянием электронов проводимости на тепловых колебаниях в простой одноатомной кубической решетке.

2. Определим изменение (возмущение) потенциальной энергии электрона в периодическом поле $V(\boldsymbol{r})$ в результате колебаний атомов на основе гипотезы деформируемых ионов Блоха (1928).

Введем представление о том, что каждая точка кристалла $\boldsymbol{r}$ «смещается» при колебаниях атомов в соответствии с формулой (III.6.27), если заменить в ней $\boldsymbol{a}_n$ на $\boldsymbol{r}$. Другими словами, «смещение» любой точки пространства между атомами рассматривается как интерполированное между значениями смещений окружающих атомов. Такая интерполяция тем более законна (в дальнейших приложениях), чем ближе по фазе колебания соседних атомов, т. е. чем длиннее акустическая волна.

В соответствии со сказанным выше, положим

$$\boldsymbol{u}=\frac{1}{\sqrt{NM}}\sum_{qj}{}'\,\boldsymbol{e}_{qj}\{a_{qj}e^{iqr}+a^*_{qj}e^{-iqr}\}, \tag{3.1}$$

где введены обозначения $\boldsymbol{e}_j(\boldsymbol{q})\equiv\boldsymbol{e}_{qj}$ и $a_j(\boldsymbol{q})\equiv a_{qj}$. Согласно (III.6.2)

$$(\boldsymbol{e}_{qj}\boldsymbol{e}_{qj'})=\delta_{jj'}. \tag{3.2}$$

Гипотеза деформируемых ионов заключается в предположении, что потенциальная энергия электрона в точке $\boldsymbol{r}$, равная до деформации кристалла $V(\boldsymbol{r})$, после деформации переносится в точку $\boldsymbol{r}+\boldsymbol{u}$, где потенциальная энергия до деформации равнялась $V(\boldsymbol{r}+\boldsymbol{u})$. Таким образом, изменение потенциальной энергии электрона в точке $\boldsymbol{r}+\boldsymbol{u}$ в линейном приближении по $\boldsymbol{u}$ равно[1])

$$\mathcal{U}\equiv\Delta V=V(\boldsymbol{r})-V(\boldsymbol{r}+\boldsymbol{u})=-(\operatorname{grad} V\boldsymbol{u}). \tag{3.3}$$

Мы разложили функцию $V(\boldsymbol{r}+\boldsymbol{u})$ в ряд по смещению $\boldsymbol{u}(u_x, u_y, u_z)$ и ограничились первым членом разложения. В (3.3) вместо $\boldsymbol{u}$ может быть подставлено выражение (3.1).

Для определения времени релаксации по формуле (2.12) надо знать вероятность перехода $W(\boldsymbol{k}, \boldsymbol{k}')$, которую мы вычислим, пользуясь теорией квантовых переходов Дирака (см. Приложение 20).

Рассмотрим систему, состоящую из электрона в периодическом поле и нормальных колебаний решетки. Волновая функция системы в нулевом приближении, т. е. без учета взаимодействия (3.3), равна произведению блоховской волновой функции электрона в идеальном кристалле (IV.3.5) и осцилляторных волновых функций (III.10.4б), т. е.

$$\Psi^{(0)}_{k,\,N_{qj}}=\psi_k(\boldsymbol{r})\prod_{qj}\psi_{N_{qj}}(Q_{qj})=\frac{1}{\sqrt{N}}u_k(\boldsymbol{r})e^{ikr}\prod\psi_{N_{qj}}(Q_{qj}), \tag{3.4}$$

где $N=G^3$—число атомов в основной области. Ввиду наличия у блоховской волновой функции множителя $N^{-1/2}$

$$\int|u_k(\boldsymbol{r})|^2d^3r_0=1, \tag{3.4а}$$

где интегрирование ведется по объему элементарной ячейки кристалла (IV.3.6).

---

[1]) Эта величина совпадает в линейном приближении по $\boldsymbol{u}$ с изменением потенциальной энергии электрона в точке $\boldsymbol{r}$.

Разложим возмущенную волновую функцию $\Psi(t)$ по полной системе [1]) невозмущенных функций (3.4):

$$\Psi(t)=\sum_{k'N'_{qj}} a\ (\boldsymbol{k}', N'_{qj}, t)\,\Psi^{(0)}_{\boldsymbol{k}'N'_{qj}}\times$$
$$\times\exp\left\{-\frac{i}{\hbar}\left[\varepsilon_{\boldsymbol{k}'}+\sum_{qj}\left(N'_{qj}+\frac{1}{2}\right)\hbar\omega_{qj}\right]t\right\}, \quad (3.5)$$

где $a(\boldsymbol{k}', N'_{qj}, t)$ — неизвестные коэффициенты разложения [2]) и

$$\varepsilon_{\boldsymbol{k}'}+\sum_{qj}\left(N'_{qj}+\frac{1}{2}\right)\hbar\omega_{qj}=\mathcal{E}_{\boldsymbol{k}',\,N'_{qj}} \quad (3.5a)$$

есть энергия невозмущенной системы, равная сумме энергий электрона и всех нормальных колебаний. Если в начальный момент $t=0$ электрон находился в состоянии $\boldsymbol{k}$, а нормальные колебания характеризовались квантовыми числами N_{qj}, то

$$a(\boldsymbol{k}', N'_{qj}, 0)=\begin{cases} 1 \text{ при } \boldsymbol{k}'=\boldsymbol{k} \text{ и } N'_{qj}=N_{qj}, \\ 0 \text{ во всех остальных случаях.} \end{cases} \quad (3.6)$$

Из общей теории имеем [3])

$$i\hbar\,\frac{da(\boldsymbol{k}', N'_{qj}, t)}{dt}=\langle\boldsymbol{k}', N'_{qj}|\,\Delta V\,|\boldsymbol{k}, N_{qj}\rangle\times$$
$$\times\exp\left\{\frac{i}{\hbar}\left[\varepsilon_{\boldsymbol{k}'}-\varepsilon_{\boldsymbol{k}}+\sum_{qj}(N'_{qj}-N_{qj})\hbar\omega_{qj}\right]t\right\}, \quad (3.7)$$

где матричный элемент перехода из состояния $\boldsymbol{k}$, N_{qj} в состояние k', N'_{qj} под действием возмущения (3.3) равен

$$\langle\boldsymbol{k}', N'_{qj}|\,\Delta V\,|\boldsymbol{k}, N_{qj}\rangle=\int\Psi^{*(0)}_{\boldsymbol{k}',\,N'_{qj}}\Delta V\Psi^{(0)}_{\boldsymbol{k},\,N_{qj}}\,d^3r\prod_{qj}dQ_{qj}. \quad (3.7a)$$

Здесь интеграл берется по трем координатам электрона d^3r, по основной области кристалла и по $3N$ нормальным координатам $\prod_{qj}dQ_{qj}$ колебаний. В экспоненте выражения (3.7) стоит разность энергий конечного и начального состояний системы.

Вычисление матричного элемента (3.7а) несколько громоздко и приведено в Приложении 21. Можно показать, что электрон взаимодействует только с одной продольной ветвью колебаний, для которой $\boldsymbol{e}_{qj}$ параллельно $\boldsymbol{q}$ [4]). Вычисление показывает, что

---

[1]) Строго говоря, полная система электронных волновых функций включает состояния, относящиеся и к другим энергетическим зонам; однако в исследуемых нами процессах рассеяния их «примесь» практически равна нулю.

[2]) Экспоненциальные множители в (3.5) выделены для удобства.

[3]) Шифф Л., Квантовая механика.— М.: ИЛ, 1957, § 29.

[4]) В дальнейшем мы для простоты опускаем индекс поляризации j.

матричный элемент (3.7а) отличен от нуля, т. е. переходы системы $\boldsymbol{k}, N_{qj} \to \boldsymbol{k}', N'_{qj}$ возможны в двух случаях:

1) волновой вектор и энергия электрона в конечном состоянии равны

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{q} \text{ и } \varepsilon_{\boldsymbol{k}'} = \varepsilon_{\boldsymbol{k}} + \hbar\omega_{\boldsymbol{q}}, \tag{3.8}$$

при этом число $\boldsymbol{q}$-фононов в конечном состоянии равно

$$N'_{\boldsymbol{q}} = N_{\boldsymbol{q}} - 1; \tag{3.8а}$$

2) волновой вектор и энергия электрона в конечном состоянии равны

$$\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{q}, \quad \varepsilon_{\boldsymbol{k}'} = \varepsilon_{\boldsymbol{k}} - \hbar\omega_{\boldsymbol{q}}; \tag{3.9}$$

при этом

$$N'_{\boldsymbol{q}} = N_{\boldsymbol{q}} + 1. \tag{3.9а}$$

Первый процесс естественно интерпретировать как поглощение электроном одного фонона. Второй — как испускание электроном фонона. При этом как при поглощении, так и при испускании фонона имеют место законы сохранения квазиимпульса и энергии. В первом случае матричный элемент (3.7а) равен [1])

$$-\frac{2i}{3}\frac{Cq}{\sqrt{N}}\sqrt{\frac{\hbar N_{\boldsymbol{q}}}{2M\omega_{\boldsymbol{q}}}}, \tag{3.10}$$

во втором случае —

$$+\frac{2i}{3}\frac{Cq}{\sqrt{N}}\sqrt{\frac{\hbar (N_{\boldsymbol{q}}+1)}{2M\omega_{\boldsymbol{q}}}}. \tag{3.10а}$$

Здесь

$$C = \frac{\hbar^2}{2m}\int |\operatorname{grad} u_{\boldsymbol{k}}|^2 d^3 r_0 \tag{3.10б}$$

и интегрирование ведется по объему элементарной ячейки кристалла. Постоянная C, имеющая размерность энергии, характеризует интенсивность взаимодействия электрона с колебаниями решетки. Для того чтобы оценить порядок величины C, положим $|\operatorname{grad} u_{\boldsymbol{k}}| \approx u_{\boldsymbol{k}}/a$, где a — постоянная решетки, тогда по порядку величины

$$C \approx \left(\frac{\hbar^2}{2ma^2}\right)\int |u_{\boldsymbol{k}}|^2 d^3 r_0 = \left(\frac{\hbar^2}{2ma^2}\right),$$

где использовано (3.4а); для $a \approx 10^{-8}$ *см* $C \approx 5$ *эв*, т. е. порядка атомной энергии. Согласно общей теории (Приложение 20) вероятность перехода в 1 *сек*

$$W(\boldsymbol{k}, \boldsymbol{k}') = W(\boldsymbol{k}, \boldsymbol{k} \pm \boldsymbol{q}) =$$
$$= \frac{2\pi}{\hbar}|\langle \boldsymbol{k}', N'_{\boldsymbol{q}}|\Delta V|\boldsymbol{k}, N_{\boldsymbol{q}}\rangle|^2 \delta(\varepsilon_{\boldsymbol{k}'} - \varepsilon_{\boldsymbol{k}} \mp \hbar\omega_{\boldsymbol{q}}). \tag{3.11}$$

[1]) Дополнительный множитель $1/\sqrt{M}$, по сравнению с (П.21.3), связан с выражением для $\boldsymbol{u}$ (3.1).

В случае поглощения фонона

$$W^+(\boldsymbol{k}, \boldsymbol{q}) = w(q) N_q \delta(\varepsilon_{k+q} - \varepsilon_k - \hbar\omega_q), \tag{3.11а}$$

а в случае испускания—

$$W^-(\boldsymbol{k}, \boldsymbol{q}) = w(q)(N_q + 1)\delta(\varepsilon_{k-q} - \varepsilon_k + \hbar\omega_q), \tag{3.11б}$$

где

$$w(q) = \frac{4\pi}{9N} \frac{C^2 q^2}{M\omega_q}. \tag{3.11в}$$

В согласии с (1.22) вероятность перехода обратно пропорциональна объему кристалла (числу атомов N в основной области). Дельта-функция $\delta(\varepsilon_{k \pm q} - \varepsilon_k \mp \hbar\omega_q)$ выражает законы сохранения энергии и волнового вектора.

Для вычисления времени релаксации τ, обусловленного рассеянием электронов на колебаниях атомной решетки, необходимо (3.11а) и (3.11б) подставить в формулу (2.12).

§ 4. Время релаксации электронов проводимости в атомном полупроводнике и металле

1. Вычислим время релаксации (длину свободного пробега) электрона проводимости, взаимодействующего с колебаниями решетки в атомном полупроводнике. Как мы сейчас покажем, рассеяние электронов происходит при этом почти упруго, так что для вычисления τ можно воспользоваться формулой (2.12).

Если энергия электрона пропорциональна k^2, т. е. $\varepsilon_k = \hbar^2 k^2/2m^*$, и частота продольных акустических волн $\omega_{q_l} = v_0 q$, где v_0 — скорость продольных волн звука, то из законов сохранения (3.8) и (3.9) следует

$$\frac{\hbar^2 (\boldsymbol{k} \pm \boldsymbol{q})^2}{2m^*} = \frac{\hbar^2 k^2}{2m^*} \pm \hbar v_0 q, \tag{4.1}$$

где верхний знак относится к случаю поглощения, а нижний — к случаю испускания фонона. Отсюда

$$q = \mp 2k \cos\vartheta \pm \frac{2m^* v_0}{\hbar}, \tag{4.1а}$$

где ϑ — угол между направлениями векторов $\boldsymbol{k}$ и $\boldsymbol{q}$.

Оценим порядок отношения второго слагаемого к первому в правой части выражения (4.1а)

$$\frac{m^* v_0}{\hbar k} = \frac{m^* v_0}{p} \approx \frac{m^* v_0}{\sqrt{m^* k_0 T}} = \sqrt{\frac{T_{\text{кр}}}{T}}, \tag{4.1б}$$

где квазиимпульс $p = \hbar k$ заменен своим средним тепловым значением при температуре T и критическая температура $T_{\text{кр}} = m^* v_0^2/k_0$. Полагая $m^* = 10^{-27}$ *г*, $v_0 = 2 \cdot 10^5$ *см/сек*, получим

$T_{кр} \approx 1°\text{K}$; поэтому при всех температурах $T \gg 1°\,\text{K}$ можно пренебречь вторым слагаемым, так что

$$q = \mp 2k\cos\vartheta. \tag{4.1в}$$

Таким образом, из законов сохранения следует, что электроны поглощают и испускают фононы с $q \approx k$.

Из выражения (4.1в) следует, что

$$q_{\min} = 0, \quad q_{\max} = 2k. \tag{4.1г}$$

Так как для электронов среднее значение k при комнатной температуре порядка 10^7 *см*, а максимальное значение волнового вектора $\boldsymbol{q}$ по теории Дебая (III.11.13) равно $q_0 = (6\pi^2/\Omega_0)^{1/3} \approx 10^8$ *см*, то $q_{\max} \ll q_0$. Электроны взаимодействуют только с длинноволновыми фононами, поэтому линейный закон дисперсии ($\omega = v_0 q$) хорошо оправдан.

Очевидно, что пренебрежение вторым слагаемым в (4.1а) равносильно пренебрежению энергией фонона в (4.1), т. е. неупругостью рассеяния электрона. При упругом рассеянии δ-функции, входящие в (3.11а) и (3.11б), имеют вид

$$\delta(\varepsilon_{k \pm q} - \varepsilon_k) = \delta\left(\pm \frac{\hbar^2 kq}{m^*}\cos\vartheta + \frac{\hbar^2 q^2}{2m^*}\right) = \frac{m^*}{\hbar^2 kq}\,\delta\left(\frac{q}{2k} \pm \cos\vartheta\right), \tag{4.2}$$

где мы воспользовались свойством $\delta(a \pm bx) = \frac{1}{b}\,\delta\left(\frac{a}{b} \pm x\right)$ [1]).

Мы можем определить теперь время релаксации τ по формуле (2.12). При рассеянии электрона на колебаниях решетки

$$\boldsymbol{k}' = \boldsymbol{k} \pm \boldsymbol{q}, \tag{4.3}$$

где верхний знак соответствует поглощению, а нижний — испусканию фонона. Из (2.12) и (4.3) имеем

$$\frac{1}{\tau} = -\sum_{\boldsymbol{q}} W^+(\boldsymbol{k}, \boldsymbol{q})\,\frac{q_\chi}{k_\chi} + \sum_{\boldsymbol{q}} W^-(\boldsymbol{k}, \boldsymbol{q})\,\frac{q_\chi}{k_\chi}, \tag{4.4}$$

где первая сумма в правой части учитывает поглощение, вторая — испускание фононов. Мы перешли в (2.12) от суммирования по $\boldsymbol{k}'$ к суммированию по $\boldsymbol{q}$ при заданном $\boldsymbol{k}$.

Заменим суммирование по $\boldsymbol{q}$ интегралом по $\boldsymbol{q}$-пространству в сферических координатах с полярной осью, направленной по $\boldsymbol{k}$:

$$\sum_{\boldsymbol{q}} \to \frac{V}{(2\pi)^3} \int_{q_{\min}}^{q_{\max}} q^2\,dq \int_0^{\pi} \sin\vartheta\,d\vartheta \int_0^{2\pi} d\varphi. \tag{4.5}$$

На рис. VIII.5 представлены векторы $\boldsymbol{k}$, $\boldsymbol{q}$ и $\boldsymbol{\chi}$ и углы между

[1]) Шифф Л., Квантовая механика, — М.: ИЛ, 1957, с. 57.

ними. Известно, что [1])

$$\cos\alpha = \cos\vartheta\cos\beta + \sin\vartheta\sin\beta\cos\varphi. \tag{4.6}$$

Так как $q_\chi = q\cos\alpha$ и $k_\chi = k\cos\beta$, то обратное время релаксации

$$\frac{1}{\tau} = -\frac{V}{(2\pi)^3}\frac{m^*}{\hbar^2 k^2}\int\limits_{q_{\min}}^{q_{\max}} q^2\,dq\int\limits_0^\pi \sin\vartheta\,d\vartheta\int\limits_0^{2\pi} d\varphi\left\{w(q)\,N_q\delta\left(\frac{q}{2k}+\cos\vartheta\right)\times\right.$$

$$\left.\times\frac{\cos\alpha}{\cos\beta} - w(q)\,(N_q+1)\,\delta\left(\frac{q}{2k}-\cos\vartheta\right)\frac{\cos\alpha}{\cos\beta}\right\}. \tag{4.7}$$

От φ зависит только $\cos\alpha$; поэтому

$$\int\limits_0^{2\pi}\cos\alpha\,d\varphi = 2\pi\cos\vartheta\cos\beta. \tag{4.8}$$

Интегрируя по $\sin\vartheta\,d\vartheta = -d\cos\vartheta$ и используя δ-функции в первом и втором слагаемых в фигурных скобках (4.7), получим

$$\frac{1}{\tau} = \frac{V}{8\pi^2}\frac{m^*}{\hbar^2 k^3}\int\limits_{q_{\min}}^{q_{\max}} w(q)(2N_q+1)\,q^3\,dq, \tag{4.9}$$

так как для процессов поглощения и испускания фононов $q_{\min}$ и $q_{\max}$ одинаковы.

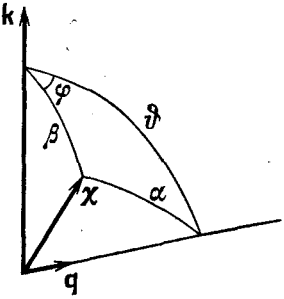


Рис. VIII. 5.

Для того чтобы использовать в (4.4) вероятности (3.11), необходимо сделать какие-то предположения о числах заполнения фононов N_q. Как мы увидим в дальнейшем, в большинстве случаев термодинамически равновесное распределение фононов, определяемое формулой Планка (III.11.10а) мало нарушается в электронных процессах переноса. В следующей главе мы увидим, что при наличии градиента температуры в электронных кинетических явлениях иногда необходимо учитывать отступление фононной функции распределения от равновесной (эффекты «увлечения фононами»).

Мы положим, таким образом,

$$N_q = \langle N_q\rangle_{\text{равн}} = \frac{1}{\exp\dfrac{\hbar v_0 q}{k_0 T} - 1}. \tag{4.10}$$

Учитывая (4.1г), видим, что показатель экспоненты

$$\frac{\hbar v_0 q}{k_0 T} \approx \frac{\hbar v_0 k}{k_0 T} = \frac{v_0 p}{k_0 T} \approx \frac{v_0\sqrt{m^* k_0 T}}{k_0 T} = \sqrt{\frac{T_{kp}}{T}} \ll 1. \tag{4.10а}$$

[1]) Кочин Н. Е., гл. I.

Поэтому (4.10) можно разложить в ряд, и мы получим

$$\langle N_q\rangle_{\text{равн}} \approx \langle N_q+1\rangle_{\text{равн}} = \frac{k_0T}{\hbar v_0 q}. \qquad (4.10б)$$

Подставляя это значение N_q и $w(q)$ из (3.11в) в (4.9), получим после элементарного интегрирования по q в пределах (4.1г)

$$\tau = \frac{9\pi}{4}\frac{Mv_0^2\hbar^3}{\Omega_0 C^2 m^* k_0 T}\frac{1}{k} = \frac{9\pi}{4\sqrt{2}}\frac{Mv_0^2\hbar^4}{\Omega_0 C^2 m^{*3/2} k_0 T}\frac{1}{\sqrt{\varepsilon}} \qquad (4.11)$$

или

$$\tau = \frac{\tau_{0k}}{k} = \frac{\tau_0}{\sqrt{\varepsilon}}, \qquad (4.11а)$$

где

$$\tau_{0k} = \frac{9\pi}{4}\frac{(Mv_0^2\hbar^3)}{\Omega_0 C^2 m^* k_0 T} \quad \text{и} \quad \tau_0 = \frac{\hbar}{\sqrt{2m^*}}\tau_{0k}.$$

Заметим, что время релаксации носителей тока в атомных полупроводниках $\tau \sim T^{-1}\varepsilon^{-1/2}$. Длина свободного пробега

$$l = v\tau = \frac{9\pi}{4}\frac{Mv_0^2\hbar^4}{\Omega_0 C^2 m^{*2} k_0 T} \qquad (4.12)$$

от энергии электрона не зависит.

2. Рассмотрим время релаксации электронов проводимости в металле. В следующей главе будет показано, что в явлениях переноса в металле практически участвуют только электроны с энергией, близкой к энергии Ферми ζ_0. Этим электронам соответствует волновой вектор

$$k(\zeta_0) = \left(\frac{1}{\hbar}\right)\sqrt{2m^*\zeta_0} \approx 10^8\ см, \qquad (4.13)$$

т. е. на порядок больший, чем k у полупроводников при комнатной температуре. Поэтому отношение (4.1б) для металлов

$$\frac{m^* v_0}{\hbar k(\zeta_0)} = \sqrt{\frac{k_0T_k}{\zeta_0}} \ll 1 \qquad (4.13а)$$

при всех температурах. Таким образом, электроны в металле, как и в полупроводнике, рассеиваются колебаниями упруго. Вероятность перехода определяется теми же выражениями (3.11а) и (3.11б), поэтому время релаксации вычисляется совершенно аналогично тому, как это сделано выше. Однако при интегрировании по q в (4.9) необходимо иметь в виду следующее. Как мы видели выше, максимальное значение волнового вектора фононов равно по Дебаю (III.11.13) $q_0 \approx 10^8\ см$, т. е. того же порядка, что и $k(\zeta_0)$. Поэтому необходимо различать два случая: 1) $k(\zeta_0) < \frac{q_0}{2}$ и 2) $k(\zeta_0) > \frac{q_0}{2}$. В первом случае, который реали-

зуется в металлах с малой концентрацией электронов проводимости (полуметаллах), интегрирование по q надо вести до $2k(\zeta_0)$, так что время релаксации равно (4.11). Во втором случае, который обычно осуществляется в хороших металлах, интегрирование по q надо вести до q_0, и в этом единственное отличие по сравнению с тем, что мы имели для полупроводников. В результате τ для металла может быть получено из выражения (4.11), если его помножить на $(2k)^4$ и разделить на q_0^4[1]). Таким образом, для металлов

$$\tau=\frac{1}{\pi^3}\frac{\Omega_0 M k_0 T_c \hbar}{m^* C^2}\left(\frac{T_c}{T}\right)k^3=\frac{2\sqrt{2}}{\pi^3}\frac{\Omega_0 M m^{*1/2} k_0 T_c}{\hbar^2 C^2}\left(\frac{T_c}{T}\right)\varepsilon^{3/2} \tag{4.14}$$

и

$$l=v\tau=\frac{4}{\pi^3}\frac{\Omega_0 M k_0 T_c}{\hbar^2 C^2}\left(\frac{T_c}{T}\right)\varepsilon^2. \tag{4.14а}$$

Здесь $T_c=\hbar v_0 q_0/k_0$ — температура Дебая (мы допустили неточность, полагая скорость продольных звуковых волн равной их средней скорости (III.9.16).

Учитывая связь между T_c и v_0, оценим порядок отношения длин свободных пробегов электрона в атомном полупроводнике (4.11а) и металле (4.14а):

$$\frac{l_{\text{п}}}{l_{\text{м}}}\approx\left(\frac{\hbar^2}{m^* a^2}\right):\varepsilon\approx\left(\frac{\hbar^2}{m^* a^2}\right):\zeta_0, \tag{4.15}$$

т. е. порядка единицы (a — постоянная решетки, равная $\Omega_0^{1/3}$). Оценим длину свободного пробега электрона $l_{\text{п}}\approx l_{\text{м}}$. Из (4.12)

$$l\approx\frac{Mv_0^2}{C}\frac{\hbar^2/m^* a^2}{C}\frac{\hbar^2/m^* a^2}{k_0 T}a. \tag{4.16}$$

Так как Mv_0^2 и $\hbar^2/m^* a^2$ имеют порядок атомной энергии, то первый и второй множители — порядка 1, третий множитель — много больше 1; поэтому $l\gg a$. В реальных атомных полупроводниках l бывает в десятки, а при низких температурах в тысячи раз больше постоянной решетки a. С точки зрения классической механики длина свободного пробега электрона в кристалле должна быть порядка постоянной решетки, так как расстояние между атомами в твердом теле порядка размеров самих атомов. Большие длины свободных пробегов электронов, вытекающие из теории и наблюдаемые на опыте, свидетельствуют о том, что движение электрона подчиняется законам квантовой механики.

---

[1]) В (4.9) как в случае полупроводника, так и в случае металла, интеграл берется от q^3.

§ 5. Теория деформационного потенциала в кубических кристаллах с простой зонной структурой

1. В последнее время получил развитие особый метод вычисления времени релаксации для электронов проводимости в атомных кристаллах, получивший название *теории деформационного потенциала* (Бардин и Шокли, 1950). Результаты, получаемые этим методом, в случае простой зонной структуры электронов совпадают с теми, какие получаются из гипотезы деформируемых ионов (§ 3) или гипотезы жестких ионов Нордгейма[1]). Однако теория деформационного потенциала обладает рядом преимуществ, главнейшее из которых — простота вывода, позволяющая обобщить теорию на более сложные случаи, например на случай сложной зонной структуры носителей тока. В этом параграфе мы рассмотрим теорию деформационного потенциала в простом кубическом одноатомном кристалле для случая простой зонной структуры, когда энергия электрона проводимости

$$\varepsilon = \hbar^2 \boldsymbol{k}^2 / 2m^*. \tag{5.1}$$

Для того чтобы более наглядно представить себе особенности метода деформационного потенциала, рассмотрим вопрос об образовании энергетических зон в кристаллах при сближении изолированных атомов. Представим себе атомы какого-либо элемента расположенными в узлах соответствующей элементу «решетки», но на больших расстояниях a друг от друга.

До тех пор, пока a много больше истинной (равновесной) постоянной решетки a_0, атомы не взаимодействуют друг с другом и их энергетические термы остаются дискретными. На рис. VIII.6, *а* видно, что при $a > 3a_0$ $3s$- и $3p$-уровни атома натрия остаются практически дискретными. При уменьшении a, т. е. сближении атомов, из-за их взаимодействия дискретные уровни атомов расширяются, образуя энергетические зоны. Как видно из рис. VIII.6, *а* при $a = 1{,}7a_0$ расширенные термы $3p$ и $3s$ начинают перекрываться, образуя при $a = 1{,}7a_0$ зону проводимости металла натрия.

Другая ситуация возникает, например, в случае кристалла алмаза. Дискретные $2s$- и $2p$-состояния изолированного атома углерода при уменьшении a расширяются и при некотором значении $a > a_0$ $2s$- и $2p$-зоны перекрываются (рис. VIII.6, *б*). Однако, как показывает расчет, при дальнейшем сближении атомов углерода происходит гибридизация s- и p-состояний (за счет образования четырех валентных связей с ближайшими соседями),

---

[1]) Пикус Г. Е. — Журнал технической физики, 1958, т. 28, с. 2390.

и при значении $a = a_0$ разрешенные зоны смешанного s—p-типа оказываются вновь разделенными запрещенной зоной.

Из рис. VIII.6, *б* видно, что при равномерном всестороннем сжатии кристалла, сопровождающемся уменьшением постоянной

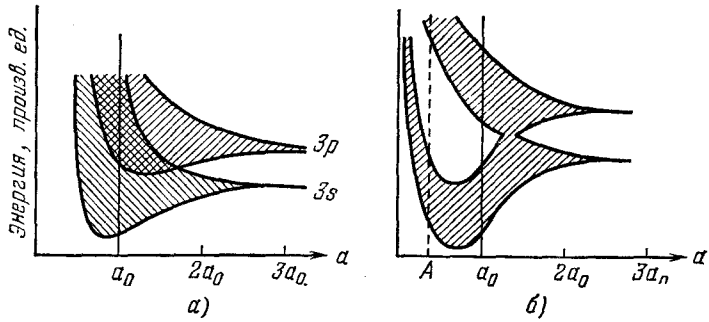


Рис. VIII. 6.

решетки a, нижний край зоны проводимости смещается вверх, а верхний край валентной зоны — вниз; в результате этого ширина запрещенной зоны увеличивается. Если бы равновесная постоянная решетки a_0 соответствовала точке A, то имела бы место иная ситуация. В этом смысле влияние сжатия и растяжения решетки может существенным образом отличаться от влияния внешнего электрического поля, которое всегда смещает край зоны проводимости и край валентной зоны в одном направлении. На рис. VIII.7 представлена картина волнообразных колебаний ширины запрещенной зоны при прохождении продольной акустической волны сжатия.

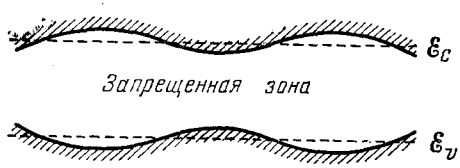


Рис. VIII. 7.

В приближении упругого континуума состояние деформированного кристалла характеризуется компонентами тензора деформации

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right), \qquad (5.2)$$

где x_i ($x_1 \equiv x$, $x_2 \equiv y$, $x_3 \equiv z$ — прямоугольные координаты точки континуума), а u_i ($i = 1, 2, 3$) — прямоугольные проекции смещения $\boldsymbol{u}(x_1, x_2, x_3)$ точки континуума x_1, x_2, x_3.

В общем случае положение нижнего края зоны проводимости $\mathcal{E}_c$ (верхнего края валентной зоны $\mathcal{E}_v$) можно рассматривать как функцию компонент тензора деформации ε_{ij}. Разлагая $\mathcal{E}_c$ в ряд

по ε_{ij}, получим

$$\mathscr{E}_c(\varepsilon_{ij}) = \mathscr{E}_c(0) + \sum_{i,j} a_{ij}\varepsilon_{ij} = \mathscr{E}_c(0) + a_{11}\varepsilon_{11} + a_{12}\varepsilon_{12} + \ldots \quad (5.3)$$

Величины a_{ij} и ε_{ij} зависят от ориентации координатных осей x_i относительно осей кристалла. Кроме того, a_{ij} зависят, конечно, и от природы кристалла. Поместим начало координат нашей прямоугольной системы в вершину куба недеформированной кристаллической ячейки и направим оси координат по его ребрам. Нетрудно показать, что для кристалла кубической симметрии недиагональные коэффициенты a_{ij} $(i \neq j)$ равны нулю. В самом деле, повернем координатную систему вокруг оси $x_3 \equiv z$ на угол $\pi/2$; тогда в новой (штрихованной) системе $x_1' = x_2$, $x_2' = -x_1$ и, следовательно, согласно (5.2) $\varepsilon_{12}' = -\varepsilon_{12}$. Коэффициенты a_{ij} при таком повороте координатной системы не меняются, так как кристалл ориентирован одинаково при обоих положениях системы, так что $a_{12}' = a_{12}$. Пусть деформация кристалла такова, что отлична от нуля только компонента тензора ε_{12}. В этом случае смещение края зоны проводимости, выраженное в обоих координатных системах, повернутых друг относительно друга, равно

$$\mathscr{E}_c(\varepsilon_{12}) - \mathscr{E}_c(0) = a_{12}\varepsilon_{12} = a_{12}'\varepsilon_{12}' = -a_{12}\varepsilon_{12},$$

откуда $a_{12} = 0$. Аналогично можно показать, что все остальные недиагональные коэффициенты a_{ij} тоже равны нулю. Так как в кубическом кристалле оси x_1, x_2, x_3 равноценны, то $a_{11} = a_{22} = a_{33} \equiv \mathscr{E}_1$ и из (5.3) имеем

$$\mathscr{E}_c(\varepsilon_{ij}) = \mathscr{E}_c(0) + \mathscr{E}_1\Delta, \quad (5.4)$$

где

$$\Delta = \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} = \operatorname{div} \boldsymbol{u} = \frac{\delta V}{V} \quad (5.4a)$$

— относительное изменение объема в данной точке [1]).

В теории деформационного потенциала Бардина и Шокли доказывается, что при рассмотрении рассеяния электронов на колебаниях решетки можно электрон описывать не блоховской волновой функцией, а плоской волной, если при этом заменить потенциал рассеяния (3.3) выражением

$$\mathscr{U} = \mathscr{E}_1\Delta = \mathscr{E}_1 \operatorname{div} \boldsymbol{u}(\boldsymbol{r}). \quad (5.5)$$

---

[1]) Зоммерфельд А. Механика деформируемых сред.—М.: ИЛ, 1954, § 1.

Мы увидим, что константа $\mathscr{E}_1$, имеющая размерность энергии, тесно связана с константой взаимодействия C, определенной в (3.10б)[1]).

2. Определим время релаксации, используя потенциал рассеяния (5.5) и описывая электрон проводимости плоской волной:

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} e^{i\boldsymbol{k}\boldsymbol{r}}, \tag{5.6}$$

где V — объем основной области кристалла.

Смещение $\boldsymbol{u}(\boldsymbol{r})$ точки $\boldsymbol{r}$ кристалла в приближении непрерывного континуума, согласно (3.1), равно

$$\boldsymbol{u}(\boldsymbol{r}) = \frac{1}{\sqrt{NM}} \sum_{q,\,j}{}' \boldsymbol{e}_{qj} \{a_{qj} e^{iqr} + a^*_{qj} e^{-iqr}\}. \tag{5.7}$$

Здесь N — число атомов основной области, $\boldsymbol{e}_{qj}$ — ортонормированные векторы поляризации, a_{qj} — комплексные нормальные координаты, гармонически зависящие от времени с частотой $\omega_{qj} = v_{0j} q$, где v_{0j} — скорость звука с поляризацией j. Отсюда относительное изменение объема

$$\Delta = \operatorname{div} \boldsymbol{u} = \frac{i}{\sqrt{NM}} \sum_{q}{}' (\boldsymbol{e}_{qj}\boldsymbol{q}) \{a_{qj} e^{iqr} - a^*_{qj} e^{-iqr}\}. \tag{5.8}$$

Так как потенциал рассеяния (5.5) пропорционален Δ, то из (5.8) следует, что электроны проводимости взаимодействуют только с продольной акустической волной, для которой $\boldsymbol{e}_{qj} \| \boldsymbol{q}$ и $\boldsymbol{e}_{qj}\boldsymbol{q} = q$. В дальнейшем мы опустим для продольной волны индекс поляризации j.

Для определения вероятности перехода электрона $W(\boldsymbol{k}, \boldsymbol{k}')$ по формуле (3.11) вычислим матричный элемент энергии возмущения (5.5) на волновых функциях электрона в приближении плоских волн (5.6), т. е. в приближении эффективной массы (5.1):

$$M_{\boldsymbol{k}\boldsymbol{k}'} = \langle \boldsymbol{k}' | \mathscr{U} | \boldsymbol{k} \rangle = \int_V \psi^*_{\boldsymbol{k}'}(\boldsymbol{r})\, \mathscr{U}(\boldsymbol{r})\, \psi_{\boldsymbol{k}}(\boldsymbol{r})\, d\tau. \tag{5.9}$$

Используя (5.5) и (5.8), получим

$$M_{\boldsymbol{k}\boldsymbol{k}'} = \frac{i\mathscr{E}_1}{\sqrt{NM}} \sum_{qj}{}' q \left\{ a_q \frac{1}{V} \int_V e^{i(\boldsymbol{k}+\boldsymbol{q}-\boldsymbol{k}')\,\boldsymbol{r}}\, d\tau - a^*_q \frac{1}{V} \int_V e^{i(\boldsymbol{k}-\boldsymbol{q}-\boldsymbol{k}')\,\boldsymbol{r}}\, d\tau \right\}. \tag{5.10}$$

---

[1]) Выражение для смещения края зоны проводимости (5.4) было одновременно и независимо использовано М. Ф. Дейгеном и С. И. Пекаром для расчета «конденсонного» состояния электрона в кубическом атомном кристалле. Титейка в 1935 г. использовал (5.5) в качестве потенциала рассеяния электронов в металле при рассчете гальваномагнитных явлений в сильных магнитных полях. Заслуга Бардина и Шокли заключается в обосновании теории деформационного потенциала в рамках метода эффективной массы.

Первый интеграл в фигурных скобках отличен от нуля и равен V только в случае

$$\boldsymbol{k}+\boldsymbol{q}=\boldsymbol{k}', \tag{5.10а}$$

второй — только в случае

$$\boldsymbol{k}-\boldsymbol{q}=\boldsymbol{k}'. \tag{5.10б}$$

Случай (5.10а) соответствует процессу, при котором электрон поглощает фонон с волновым вектором $\boldsymbol{q}$ и энергией $\hbar\omega_q$, а случай (5.10б) соответствует процессу испускания фонона.

В обоих случаях

$$|M_{\boldsymbol{k}\boldsymbol{k}'}|^2=\frac{\mathscr{E}_1^2}{NM}q^2|a_q|^2, \tag{5.11}$$

так как $|a_q|^2=|a_q|^2$.

Из (III.6.27) следует, что энергия одного нормального колебания, соответствующего продольной акустической волне, равна $2\omega_q^2|a_q|^2=2v_0^2q^2|a_q|^2$, где v_0 — скорость продольных волн звука. В состоянии статистического равновесия (при температуре выше дебаевской)

$$2v_0^2q^2|a_q|^2=k_0T, \tag{5.12}$$

откуда

$$|a_q|^2=k_0T/2v_0^2q^2. \tag{5.12а}$$

Из (5.11) и (5.12а) следует, что

$$|M_{\boldsymbol{k}\boldsymbol{k}'}|^2=\mathscr{E}_1^2k_0T/2NMv_0^2. \tag{5.13}$$

Из (3.10) и (3.10а) следует, что для температур выше дебаевской

$$\langle\boldsymbol{k}', N'_{qj}|\Delta V|\boldsymbol{k}, N_{qj}\rangle=\frac{4}{9}\frac{c^2q^2}{N}\frac{\hbar}{2M\omega_q}\frac{k_0T}{\hbar\omega_q}=\frac{4}{9}\frac{C^2k_0T}{2NMv_0^2}. \tag{5.14}$$

Здесь учтено, что $N_q\approx N_q+1=k_0T/\hbar\omega_q$ и $\omega_q=v_0q$. Если считать квадраты модулей матричных элементов (5.13) и (5.14) одинаковыми, то

$$\mathscr{E}_1={}^2/_3C, \tag{5.15}$$

где константа взаимодействия C определяется выражением (3.10б).

Из (5.14) непосредственно вытекает выражение для вероятности перехода (3.11), которое позволяет нам, подобно тому как это было сделано в § 4, вычислить время релаксации электрона τ. Очевидно, что полученное таким образом время релаксации совпадает с выражением (4.11), в котором надо только заменить C на ${}^3/_2\mathscr{E}_1$.

Конечно, нельзя рассматривать приведенные выше рассуждения как прямое доказательство соотношения (5.15), так как оно основано на сравнении конечных результатов вычислений, исходящих из разных физических предпосылок. В Приложении 22 дается прямое доказательство соотношения (5.15), основанное на квантовомеханической теории возмущений деформированного кристалла.

§ 6. Рассеяние электронов проводимости в ионных кристаллах на колебаниях решетки

1. Электрон проводимости (дырка) в ионном кристалле (если он не пьезоэлектрик) гораздо сильнее взаимодействует с оптическими колебаниями, чем с акустическими. Это связано с тем, что в ионном кристалле при оптических колебаниях в каждой кристаллической ячейке возникает электрический дипольный момент, с которым сильно взаимодействует электрон (дырка) проводимости.

Исследуя поляроны слабой связи (гл. V, § 4, п. 3), мы определили матричные элементы взаимодействия электрона с длинноволновыми оптическими колебаниями ионного кристалла в континуальном приближении. Мы видели, что в первом приближении теории возмущений поправка к энергии электрона равна нулю. Вычислим в этом же приближении рассеяние электрона, связанное с рассеянием на оптических колебаниях ионного кристалла. Из (V.4.46) и (3.11) следует, что вероятность перехода электрона $\boldsymbol{k} \to \boldsymbol{k}'$, связанная с поглощением или испусканием оптического фонона $\hbar\omega_l$, равна

$$W(\boldsymbol{k}, \boldsymbol{k}') = W(\boldsymbol{k}, \boldsymbol{k} \pm \boldsymbol{q}) = \frac{2\pi}{\hbar} |\langle N'_q, \boldsymbol{k}' | -e\Phi | N_q, \boldsymbol{k}\rangle|^2 \delta(\varepsilon_{\boldsymbol{k}'} - \varepsilon_{\boldsymbol{k}} \mp \hbar\omega_l) = w(q) \begin{Bmatrix} N_q \\ N_q + 1 \end{Bmatrix} \delta(\varepsilon_{\boldsymbol{k}\pm\boldsymbol{q}} - \varepsilon_{\boldsymbol{k}} \mp \hbar\omega_l), \quad (6.1)$$

где

$$w(q) = \frac{4\pi^2 e^2 \omega_l}{\varepsilon^*} \frac{1}{q^2}. \quad (6.1a)$$

Здесь верхние знаки (и верхняя строка в фигурной скобке) соответствуют поглощению, а нижние — испусканию фонона. Закон сохранения энергии выражается δ-функцией в (6.1); если электрон (дырка) проводимости характеризуется эффективной массой m, то закон сохранения энергии имеет вид

$$\frac{\hbar^2(\boldsymbol{k} \pm \boldsymbol{q})^2}{2m^*} = \frac{\hbar^2 k^2}{2m^*} \pm \hbar\omega_l,$$

или

$$\frac{\hbar^2 q^2}{2m^*} \pm \frac{\hbar^2 kq}{m^*} \cos\vartheta \mp \hbar\omega_l = 0. \quad (6.2)$$

Здесь ϑ — угол между волновыми векторами $\boldsymbol{k}$ и $\boldsymbol{q}$. Решая квадратное уравнение (6.2) относительно q, получим два корня

$$q_1 = -k\cos\vartheta \pm \sqrt{k^2\cos^2\vartheta + k_0}, \quad q_2 = k\cos\vartheta \pm \sqrt{k^2\cos^2\vartheta - k_0}, \quad (6.3)$$

где

$$\hbar^2 k_0^2/2m^* = \hbar\omega_l. \tag{6.3a}$$

Рассмотрим теперь разные температурные интервалы.

А. Высокие температуры: $k_0 T \gg \hbar\omega_l$ или $k \gg k_0$. В этом случае можно пренебречь k_0^2 под знаками корня в (6.3) и как в случае поглощения, так и испускания фононов

$$q_{\min} = 0, \quad q_{\max} = 2k. \tag{6.4}$$

Далее, при $k_0 T \gg \hbar\omega_l$

$$N_q = \langle N_q \rangle_{\text{равн}} = \frac{k_0 T}{\hbar\omega_l} \gg 1. \tag{6.5}$$

Так как при $k_0 T \gg \hbar\omega_l$ рассеяние носит упругий характер, то время релаксации определяется по формуле (4.9)

$$\frac{1}{\tau} = \frac{1}{8\pi^2} \frac{m^*}{\hbar^2 k^3} \int_{q_{\min}}^{q_{\max}} w(q)\,(2N_q + 1)\, q^3\, dq. \tag{6.6}$$

Подставляя сюда (6.1а), (6.4) и (6.5), получим для времени релаксации электронов (дырок) проводимости в ионном кристалле при высоких температурах

$$\tau = \frac{\sqrt{2}}{2} \frac{\hbar^2 \varepsilon^*}{e^2 (m^*)^{1/2} k_0 T}\, \varepsilon^{1/2}, \tag{6.7}$$

таким образом, $\tau \backsim \sqrt{\varepsilon}/T$. Длина свободного пробега электрона тепловой энергии $\varepsilon = m^* v^2/2 = {}^3/_2 k_0 T$ равна

$$l = \tau v = \frac{3}{2} \varepsilon^* \frac{\hbar^2}{m^* e^2}. \tag{6.7a}$$

Если эффективная масса электрона $m^* = m$ — массе свободного электрона, то $\hbar^2/me^2$ — радиус боровской орбиты и длина свободного пробега l порядка постоянной решетки: в этом случае выражение (6.7) неприменимо, оно применимо только в случае большего отношения ε^*/m^*.

Б. Низкие температуры: $k_0 T \ll \hbar\omega_l$ или $k \ll k_0$. В этом случае практически имеют место только процессы поглощения фононов и, согласно (6.3),

$$q_{\min} = \sqrt{k^2 + k_0^2} - k \quad (\vartheta = 0), \quad q_{\max} = \sqrt{k^2 + k_0^2} + k \quad (\vartheta = \pi). \tag{6.8}$$

При низких температурах рассеяние электронов неупруго, поэтому вводить время релаксации по формуле (4.9), вообще говоря, нельзя. Однако, как было показано (Б. И. Давыдов и И. М. Шмушкевич, 1940), и в случае низких температур время релаксации может быть введено посредством некоторой корректной схемы расчета.

Качественно это можно понять следующим образом. При низких температурах, когда $k_0T \ll \hbar\omega_l$, подавляющее большинство электронов могут только поглощать фононы. При таком поглощении фонона электрон переходит в энергетический интервал от $\hbar\omega_l$ до $2\hbar\omega_l$. Такой электрон почти мгновенно испустит фонон, так как отношение вероятности испускания к вероятности поглощения, согласно (6.1), равно $\frac{N_q+1}{N_q} \approx \exp\frac{\hbar\omega_l}{k_0T} \gg 1$. В результате такого поглощения и почти мгновенного испускания фонона энергия электрона изменится очень мало (только за счет зависимости ω_l от q), а его волновой вектор — значительно. Это позволяет в некотором смысле рассматривать рассеяние электрона как упругое и ввести время релаксации

Вычислим время релаксации τ по формуле (4.4), учитывая в ней справа только первую сумму, соответствующую поглощению фонона.

Дельта-функция, входящая в $W^{+}_{\boldsymbol{k},\boldsymbol{k}'}$ имеет вид

$$\delta\left(\frac{\hbar^2q^2}{2m^*}-\frac{\hbar^2k_0^2}{2m^*}+\frac{\hbar^2kq}{m^*}\cos\vartheta\right)=\frac{m^*}{\hbar^2kq}\,\delta\left(\frac{q^2-k_0^2}{2kq}+\cos\vartheta\right). \qquad (6.9)$$

Проводя выкладки, аналогичные тем, которые привели к выражению (4.9), получим

$$\frac{1}{\tau}=\frac{1}{8\pi^2}\frac{m^*}{\hbar^2k^3}\int\limits_{q_{\min}}^{q_{\max}} w(q)\,N_q\,\frac{q^2-k_0^2}{q}\,q^2\,dq. \qquad (6.10)$$

Подставляя сюда (6.1а) и (6.8) и учитывая, что при низких температурах

$$N_q=\langle N_q\rangle_{\text{равн}}=\frac{1}{\exp\frac{\hbar\omega_l}{k_0T}-1}\approx\exp\left(-\frac{\hbar\omega_l}{k_0T}\right), \qquad (6.11)$$

получим

$$\frac{1}{\tau}=\frac{e^2m^*\omega_l\exp\left(-\frac{\hbar\omega_l}{k_0T}\right)}{\hbar^2\varepsilon^*k^3}\left\{2k\sqrt{k^2+k_0^2}-k_0\lg\frac{\sqrt{k^2+k_0^2}+k}{\sqrt{k^2+k_0^2}-k}\right\}. \qquad (6.12)$$

Вынося из фигурной скобки k_0^3 и разлагая оставшееся выражение в ряд по отношению $k/k_0 \ll 1$, убедимся, что первый отличный от нуля член порядка $(k/k_0)^3$, так что фигурная скобка равна

$$^4/_3\,k^3/k_0. \qquad (6.12а)$$

Из (6.12) и (6.12а) получим для времени релаксации электрона (дырки) проводимости в ионном кристалле при низких

температурах:

$$\tau = \frac{3\sqrt{2}}{2} \frac{\hbar^2 \varepsilon^* \exp \frac{\hbar\omega_l}{k_0 T}}{e^2 (m^*)^{1/2} (\hbar\omega_l)^{1/2}}, \tag{6.13}$$

таким образом, $\tau \backsim \exp(\hbar\omega_l/k_0 T)$ и не зависит от энергии электрона. Длина свободного пробега электрона тепловой энергии $\varepsilon = (mv^2/2) = {}^3/_2 k_0 T$ равна

$$l = \tau v = \frac{3\sqrt{6}}{2} \varepsilon^* \left(\frac{k_0 T}{\hbar\omega_l}\right)^{1/2} \exp \frac{\hbar\omega_l}{k_0 T} \frac{\hbar^2}{m^* e^2}. \tag{6.13а}$$

Здесь $\hbar^2/m^*e$ — радиус «эффективной» боровской орбиты; благодаря большому множителю $\exp(\hbar\omega_l/k_0 T)$ l много больше постоянной решетки и выражение для времени релаксации τ (6.13), вообще говоря, применимо при низких температурах.

2. В пьезоэлектрических полупроводниках (гл. III, § 1, п. 3), например в цинковой обманке ZnS рассеяние электронов проводимости на акустических колебаниях может быть сравнимо с рассеянием на оптических колебаниях, рассмотренных в предыдущем пункте этого параграфа. Это связано с тем, что в пьезоэлектриках длинные акустические волны (вызывающие упругие напряжения) связаны с электрической поляризацией кристалла.

Впервые пьезоэлектрическое рассеяние на акустических колебаниях было рассмотрено Мейиром и Полдером [1]). Мы изложим некоторые их результаты в упрощенном виде, достаточном для определения зависимости времени релаксации электрона τ от его энергии $\varepsilon = \hbar^2 k^2/2m$ и температуры T.

Пьезоэлектрики — ионные решетки, не обладающие центром инверсии (симметрии). Смещения положительных и отрицательных ионов ($k = 1, 2$) в этом случае, как следует из (III.6.4а), в континуальном приближении равны

$$\boldsymbol{u}^k(\boldsymbol{r}) = \frac{1}{\sqrt{N m_k}} \sum_{q,\, j}{}' \{\boldsymbol{e}_{jk}(\boldsymbol{q})\, a_j(\boldsymbol{q})\, e^{i\boldsymbol{q}\boldsymbol{r}} + \boldsymbol{e}^*_{jk}(\boldsymbol{q})\, a_j^*(\boldsymbol{q})\, e^{-i\boldsymbol{q}\boldsymbol{r}}\}, \tag{6.14}$$

где суммирование по $\boldsymbol{q}$ ведется по половине зоны Бриллюэна. Дипольный момент, возникающий при колебаниях в элементарной ячейке объема Ω_0, равен геометрической сумме смещений ионов, умноженных на их эффективные заряды $\pm e^*$.

Как мы увидим ниже, нас будут интересовать только длинные акустические волны, для которых можно считать вектор $\boldsymbol{q}$ одинаковым для обоих ионов ячейки. Вектор поляризации $\boldsymbol{P}(\boldsymbol{r})$, равный дипольному моменту единицы объема, может быть записан в виде

$$\boldsymbol{P}(\boldsymbol{r}) = \frac{e^*}{\Omega_0} \frac{1}{\sqrt{N}} \sum_{q,\, j}{}' \left(\frac{\boldsymbol{e}_{j1}(q)}{\sqrt{m_1}} - \frac{\boldsymbol{e}_{j2}(q)}{\sqrt{m_2}}\right) a_j(\boldsymbol{q})\, e^{i\boldsymbol{q}\boldsymbol{r}} + \text{к. с}, \tag{6.15}$$

[1]) Meijer H., Polder D.— Physica, 1953, v. 19, p. 255.

где к.с. означает величину комплексно-сопряженную первому слагаемому.

Поляризация (6.15) обусловливает распределенный в пространстве связанный заряд $\rho = -\operatorname{div} \boldsymbol{P}(\boldsymbol{r})$ [1]), являющийся источником электрического поля, потенциал которого Φ удовлетворяет уравнению Пуассона

$$\nabla^2\Phi = -4\pi\rho = 4\pi \operatorname{div} \boldsymbol{P}(\boldsymbol{r}) = \\ = i\frac{4\pi e^*}{\Omega_0}\frac{1}{\sqrt{N}}\sum_{qj}{}' q\left(\frac{e_{j1}}{\sqrt{m_1}} - \frac{e_{j2}}{\sqrt{m_2}}\right) a_j(\boldsymbol{q})\, e^{i\boldsymbol{qr}} + \text{к. с.} \qquad (6.16)$$

Поступая совершенно аналогично тому, как мы решали уравнение (V.4.41), получим

$$\Phi = -i\frac{4\pi e^*}{\Omega_0\sqrt{N}}\sum_{qj}{}' (\boldsymbol{q}_1\boldsymbol{h}_j)\, a_j(\boldsymbol{q})\, e^{i\boldsymbol{qr}} + \text{к. с.}, \qquad (6.17)$$

где единичный вектор $\boldsymbol{q}_1 = \boldsymbol{q}/q$ и $\boldsymbol{h}_j(\boldsymbol{q}) = \frac{1}{q}\left(\frac{e_{j1}(\boldsymbol{q})}{\sqrt{m_1}} - \frac{e_{j2}(\boldsymbol{q})}{\sqrt{m_2}}\right)$.

Выражение (6.17) отличается от (V.4.42) тем, что в случае длинноволновых оптических колебаний ионного кристалла с электроном проводимости взаимодействует только продольная волна, в то время как в случае акустических колебаний пьезоэлектрика с электроном взаимодействуют, вообще говоря, все три ветви колебаний ($h_j \neq 0$ для всех j).

Энергия возмущения, соответствующая потенциалу (6.17), равна $-e\Phi$.

Как мы сейчас покажем, первой сумме в (6.17), содержащей множитель $a_j(\boldsymbol{q})\exp(i\boldsymbol{qr})$, соответствуют процессы поглощения электроном проводимости фонона $\hbar\omega_{qj}$, поэтому, согласно (3.11), вероятность поглощения фонона $\hbar\omega_{qj}$ равна

$$W_j^+(\boldsymbol{k}, \boldsymbol{k}') = \\ = \frac{2\pi}{\hbar}\left|\langle \boldsymbol{k}', N'_{qj}| -e\Phi |\boldsymbol{k}, N_{qj}\rangle\right|^2 \delta(\varepsilon_{\boldsymbol{k}'} - \varepsilon_{\boldsymbol{k}} - \hbar\omega_{qj}) = \\ = \frac{32\pi^3 e^{*2}e^2}{\hbar\Omega_0^2 N}\left|\left\langle \boldsymbol{k}', N'_{qj}\left|\sum_{\boldsymbol{q}}{}' (\boldsymbol{q}_1\boldsymbol{h}_j)\, a_j(\boldsymbol{q})\, e^{i\boldsymbol{qr}}\right|\boldsymbol{k}, N_{qj}\right\rangle\right|^2 \times \\ \times \delta(\varepsilon_{\boldsymbol{k}'} - \varepsilon_{\boldsymbol{k}} - \hbar\omega_{qj}). \qquad (6.18)$$

В качестве волновых функций электрона, как и в (V.4.43) используем плоские волны, тогда

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{q}, \qquad (6.19)$$

что соответствует процессу поглощения электроном фонона с волновым вектором $\boldsymbol{q}$; в согласии с этим матричные элементы

[1]) Тамм И. Е., § 21.

$a_j(\boldsymbol{q})$ отличны от нуля только для $N'_{qj}=N_{qj}-1$ и равны (V.4.45):

$$\langle N'_{qj}\,|\,a_j(\boldsymbol{q})\,|\,N_{qj}\rangle=\sqrt{\frac{\hbar N_{qj}}{2\omega_{qj}}}\,. \tag{6.20}$$

к. с.-член в (6.17), содержащий множитель $a_j^*(\boldsymbol{q})\exp(-i\boldsymbol{qr})$, соответствует процессу испускания электроном фонона, для которого

$$\boldsymbol{k}'=\boldsymbol{k}-\boldsymbol{q}, \tag{6.19а}$$

и

$$\langle N'_{qj}\,|\,a_j^*(\boldsymbol{q})\,|\,N_{qj}\rangle=\sqrt{\frac{\hbar(N_{qj}+1)}{2\omega_{qj}}}, \tag{6.20а}$$

если $N'_{qj}=N_{qj}+1$.

Для длинноволновых акустических фононов $\omega_{qj}=v_{qj}q$, где v_{qj}—скорость звука, зависящая от направления вектора $\boldsymbol{q}$. Взаимодействие электронов с фононами (для всех $T\gg 1^\circ$ K происходит упруго (см. § 4), δ-функция, входящая в (6.18), равна (4.2) и

$$\langle N_{qj}\rangle_{\text{равн}}\approx\langle N_{qj}+1\rangle_{\text{равн}}=k_0T/\hbar v_{qj}q, \tag{6.21}$$

как следует из (4.10б).

В этом приближении вероятность, связанная с испусканием фонона W_j^-, равна вероятности поглощения фонона, так что полная вероятность взаимодействия с фононом $W_j=W_j^++W_j^-=2W_j^+$. Для времени релаксации, связанной с взаимодействием с j-й ветвью, из (4.2)—(4.7), (6.18), (6.20) и (6.21), получим

$$\frac{1}{\tau_j}=-\sum_{\boldsymbol{q}}2W_j^+(\boldsymbol{k},\,\boldsymbol{k}+\boldsymbol{q})\frac{q_\chi}{k_\chi}=-\frac{e^{*2}e^2m^*k_0T}{\hbar^3\Omega_0k^2}\times$$

$$\times\int\limits_{q_{\min}}^{q_{\max}}dq\int\limits_0^\pi\sin\vartheta\,d\vartheta\int\limits_0^{2\pi}d\varphi\left\{\left[\frac{|\boldsymbol{q}_1\boldsymbol{h}_j|^2}{v_{qj}^2}\right]\frac{\cos\alpha}{\cos\beta}\,\delta\left(\frac{q}{2k}+\cos\vartheta\right)\right\}, \tag{6.22}$$

при этом мы учли, что $V/N=\Omega_0$.

Общие соображения и конкретные расчеты, выполненные Мейиром и Полдером[1]), показывают, что множитель $\left[\frac{|\boldsymbol{q}_1\boldsymbol{h}_j|^2}{v_{qj}^2}\right]$ не зависит от абсолютной величины вектора $\boldsymbol{q}$, а только от его направления, т. е. от углов ϑ и φ. Для упрощения вынесем среднее значение $\left[\frac{|\boldsymbol{q}_1\boldsymbol{h}_j|^2}{v_q^2}\right]$ по направлениям из-под знака

[1]) См. сноску на с. 484.

интеграла; тогда интегрирование по φ дает

$$\int_0^{2\pi} \frac{\cos\alpha}{\cos\beta}\, d\varphi = 2\pi\cos\vartheta,$$

а интегрирование по ϑ, как и в случае (4.7), приводит к множителю $q/2k$. В результате из (6.22) получим

$$\frac{1}{\tau_j} = \frac{\pi e^{*2}e^2 m^* k_0 T}{\hbar^3\Omega_0 k^3}\left[\frac{|\boldsymbol{q}_1\boldsymbol{h}_j|^2}{v_{qj}^2}\right]\int_0^{2k} q\,dq = \frac{2\pi e^{*2}e^2 m^*}{\hbar^3\Omega_0}\left[\frac{|\boldsymbol{q}_1\boldsymbol{h}_j|^2}{v_{qj}^2}\right]\frac{k_0 T}{k}. \quad (6.23)$$

До тех пор, пока не вычислен множитель $\left[\frac{|\boldsymbol{q}_1\boldsymbol{h}_j|^2}{v_{qj}^2}\right]$, выражение для времени релаксации (6.23) интересно только своей зависимостью от температуры T и энергии электрона $\varepsilon = \hbar^2k^2/2m^*$. Очевидно, что для всех трех акустических ветвей эта зависимость одинакова. Так как $1/\tau_{ак} = \sum_{j=1}^{3} 1/\tau_j$, где $\tau_{ак}$ — время релаксации взаимодействия со всеми акустическими ветвями, то

$$\frac{1}{\tau_{ак}} \backsim \frac{T}{k} \backsim \frac{T}{\sqrt{\varepsilon}}. \quad (6.24)$$

Мы видим, что в пьезоэлектрике зависимость времени релаксации $\tau_{ак}$ от температуры T и энергии электрона ε такая же, как для взаимодействия электрона с оптическими колебаниями в ионных кристаллах типа NaCl при высоких температурах (6.7). В упомянутой выше статье Мейир и Полдер пытаются вычислить $\boldsymbol{q}_1\boldsymbol{h}_j$ и v_{qj}, используя теорию упругости для пьезоэлектриков. Им удается выразить значение этих величин через упругие константы и пьезоэлектрический модуль только для основных направлений в кубическом кристалле (100), (110) и (111); эти значения они затем используют для приближенного определения времени релаксации $\tau_{ак}$. Подставляя в полученное выражение численные значения упругих констант и пьезоэлектрического модуля для ZnS, они получают

$$\frac{1}{\tau_{ак}} = 5{,}1\cdot 10^{10}\sqrt{\frac{T}{x}}, \quad (6.25)$$

где $x = \varepsilon/k_0T$.

Обратное время релаксации для взаимодействия электрона с оптическими колебаниями при низких температурах (6.13) для того же вещества равно

$$\frac{1}{\tau_{оп}} = 5{,}2\cdot 10^{13} e^{-538/T}. \quad (6.26)$$

В табл. VIII.1 приведены значения обоих обратных времен релаксации для электрона с энергией $\varepsilon = 3k_0T/2$ при разных

Таблица VIII.1

$T°$ K	$\frac{1}{\tau_{ак}} \cdot 10^{11}$, $сек^{-1}$	$\frac{1}{\tau_{оп}} \cdot 10^{11}$, $сек^{-1}$
250	6,6	60
200	5,9	53
150	5,1	14
125	4,5	7,0
100	4,2	2,4
75	3,6	0,39

температурах. Мы видим, что $1/\tau_{ак}$ становится того же порядка, что и $1/\tau_{оп}$ при температуре 125° K, и при понижении температуры рассеяние на пьезоакустических колебаниях становится все более существенным.

§ 7. Рассеяние электронов проводимости на заряженных и нейтральных атомах примесей

1. В конце § 1 мы рассмотрели с точки зрения классической механики задачу о рассеянии носителей тока на ионах примеси. В качестве рассеивающего потенциала был взят кулоновский потенциал в среде с диэлектрической постоянной ε_0 (1.17). Для того чтобы обеспечить сходимость интеграла (1.23), определяющего время релаксации, интегрирование по углу рассеяния θ велось от θ_{min}, соответствующего максимальному прицельному расстоянию $b_{max} = {}^1/_2 n_I^{-1/3}$, где n_I — концентрация ионов примеси. Было отмечено, что такое ограничение прицельного расстояния величиной, равной половине среднего расстояния между соседними ионами примеси, соответствует представлению об экранировании зарядов ионов противоположными зарядами носителей тока. В настоящем параграфе мы учтем это экранирование более последовательным образом.

Пусть n — средняя (равномерная) концентрация свободных электронов в кристалле, а $n'(\boldsymbol{r})$ — их концентрация в поле примесного иона. Если φ — потенциал электрического поля, создаваемый точечным положительным зарядом иона $+e$, помещенного в начале координат, и отрицательным зарядом избытка электронов $-e(n'-n)$, то уравнение Пуассона имеет вид

$$\nabla^2 \varphi = \frac{4\pi e}{\varepsilon_0}(n'-n), \qquad (7.1)$$

где ε_0 — диэлектрическая постоянная и

$$\lim_{r \to 0} \varphi = e/\varepsilon_0 r, \qquad (7.1a)$$

так как в начале координат помещен заряд $+e$.

Согласно (VI.2.5)

$$n=\frac{(2m^*k_0T)^{3/2}}{2\pi^2\hbar^3}\mathscr{F}_{1/2}(z), \tag{7.2}$$

где $\mathscr{F}_{1/2}(z)$ — интеграл Ферми (VI.2.6) индекса $^1/_2$, а $z=\zeta/k_0T$ (ζ — химический потенциал). Если поле φ меняется плавно (гл. V, § 1), то потенциальная энергия электрона ($-e\varphi$) добавляется к его энергии $\varepsilon=\hbar^2k^2/2m^*$ в функции распределения, так что

$$f_0=\left[\exp\left(\frac{\varepsilon-e\varphi-\zeta}{k_0T}\right)+1\right]^{-1}. \tag{7.3}$$

Последнее эквивалентно тому, что мы ζ заменяем $\zeta+e\varphi$, поэтому

$$n'=\frac{(2m^*k_0T)^{3/2}}{2\pi^2\hbar^3}\mathscr{F}_{1/2}(z+u), \tag{7.4}$$

где $u=e\varphi/k_0T$.

Если $e\varphi\ll\zeta$ (что иногда выполняется плохо), то можно разложить (7.4) в ряд по степеням u и ограничиться первым членом, тогда

$$n'=n+\frac{(2m^*k_0T)^{3/2}}{2\pi^2\hbar^3}\mathscr{F}'_{1/2}(z)\,u, \tag{7.5}$$

где

$$\mathscr{F}'_{1/2}(z)\equiv\frac{\partial\mathscr{F}_{1/2}(z)}{\partial z}.$$

Подставляя это в (7.1), получим

$$\nabla^2\varphi=q^2\varphi, \tag{7.6}$$

где

$$q^2=\frac{4\sqrt{2}}{\pi}\frac{e^2m^{*3/2}(k_0T)^{1/2}}{\varepsilon_0\hbar^3}\mathscr{F}'_{1/2}\left(\frac{\zeta}{k_0T}\right). \tag{7.6а}$$

Решение уравнения (7.6), обладающего сферической симметрией и удовлетворяющего условию (7.1а), имеет вид

$$\varphi=\frac{e}{\varepsilon_0 r}e^{-qr}. \tag{7.6б}$$

В этом можно непосредственно убедиться, подставляя (7.6б) в (7.6) и учитывая, что в сферически симметричном случае

$$\nabla^2\varphi=\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\varphi}{dr}\right).$$

При $r\approx 1/q=r_0$ потенциал φ, согласно (7.6б), заметно (в e раз) уменьшается, поэтому r_0 называется *радиусом экранирования*.

Определим r_0 в случаях слабого и сильного вырождения свободных электронов.

А. Невырожденные электроны (гл. VI, § 2.3).
В этом случае

$$\mathscr{F}'_{1/2}(z) = \frac{d}{dz}\int_0^\infty e^{z-x}x^{1/2}\,dx = \frac{\sqrt{\pi}}{2}e^z, \tag{7.7}$$

$$e^z = e^{\zeta/k_0T} = \frac{4\pi^3 n\hbar^3}{(2\pi m^* k_0 T)^{3/2}}. \tag{7.7a}$$

Из (7.6а) и (7.7) следует, что радиус экранирования

$$r_0 = \frac{1}{q} = \sqrt{\frac{\varepsilon_0 k_0 T}{4\pi e^2 n}} \tag{7.8}$$

совпадает с дебаевским радиусом (VI.7.4а).
Записывая (7.8) в виде

$$\frac{r_0}{d} = \sqrt{\frac{k_0T}{4\pi e^2/\varepsilon_0 d}}, \tag{7.8a}$$

где $d = n^{-1/3}$ — среднее расстояние между электронами, видим, что по порядку величины отношение r_0/d равно квадратному корню из отношения тепловой энергии электрона к кулоновской энергии взаимодействия пары электронов, разделенных расстоянием d.

Б. Вырожденные электроны (гл. VI, § 2, п. 2).
В этом случае

$$\mathscr{F}'_{1/2}(z) = z^{1/2} = \frac{\pi\hbar}{(2m^*k_0T)^{1/2}}\left(\frac{3n}{\pi}\right)^{1/3}. \tag{7.9}$$

Подставляя (7.9) в (7.6а), получим для радиуса экранирования

$$r_0 = \frac{1}{q} = \left[\frac{\varepsilon_0\hbar^2}{4m^*e^2}\left(\frac{\pi}{3n}\right)^{1/3}\right]^{1/2}. \tag{7.10}$$

В этом случае радиус экранирования по порядку величины определяется из условия

$$\frac{\hbar^2}{2m^*r_0^2} \approx \frac{e^2}{\varepsilon_0 d}, \tag{7.11}$$

т. е. квантовая энергия электрона в области с линейными размерами r_0 порядка энергии кулоновского взаимодействия электронов, находящихся на среднем расстоянии друг от друга d.

2. Согласно (1.23а) время релаксации

$$\frac{1}{\tau} = 2\pi v n_I \int \sigma(\theta)(1-\cos\theta)\sin\theta\,d\theta, \tag{7.12}$$

где v — скорость электрона, $\sigma(\theta)$ — дифференциальное сечение рассеяния и n_I — концентрация ионов примеси.

Для определения $\sigma(\theta)$ мы воспользуемся приближением Борна [1]) квантовомеханической теории рассеяния. В методе Борна элект-

[1]) Блохинцев Д. И., § 78.

рон рассматривается как свободный и рассеяние трактуется как малое возмущение. Это означает, что метод Борна применим только к достаточно быстрым электронам. При рассеянии на потенциале вида (7.6б) волновой вектор свободного электрона k меняется на величину порядка q, поэтому метод Борна применим только в случае $q \ll k$. Легко видеть, что для электронов с энергией $\hbar^2k^2/2m^* \approx k_0T$ и q, определяемым равенствами (7.8) или (7.10), метод Борна применим только при достаточно высоких температурах.

Нетрудно показать, что в борновском приближении сечение рассеяния на потенциале (7.6б) равно [1])

$$\sigma(\theta) = \left[\frac{e^2/\varepsilon_0}{m^*v^2(1-\cos\theta)+\dfrac{\hbar^2q^2}{2m^*}}\right]^2. \tag{7.13}$$

При $q=0$ потенциал (7.6б) переходит в чисто кулоновский и сечение (7.13) — в формулу Резерфорда (1.20). Подставляя (7.13) в (7.12), вводя в интеграле новую переменную интегрирования $t=\cos\theta$ ($dt=-\sin\theta\, d\theta$), получим в результате элементарного интегрирования

$$\tau = \frac{\varepsilon_0^2 m^{*2} v^3}{2\pi e^4 n_I \Phi(\eta)} = \frac{\sqrt{2m^*}\,\varepsilon_0^2 \varepsilon^{3/2}}{\pi e^4 n_I \Phi(\eta)}, \tag{7.14}$$

где энергия $\varepsilon = m^*v^2/2$, а

$$\Phi(\eta) = \ln(1+\eta) - \frac{\eta}{1+\eta} \tag{7.14а}$$

— медленно меняющаяся функция аргумента

$$\eta = 4m^{*2}v^2/\hbar^2q^2 = 8m^*\varepsilon/\hbar^2q^2. \tag{7.14б}$$

Ввиду медленного изменения функции $\Phi(\eta)$ можно с достаточной степенью точности считать, что $\tau \sim \varepsilon^{3/2}$, как это имело место и в формуле Конвелл — Вайскопфа (1.25).

3. Рассеяние носителей тока на нейтральных атомах примеси может приближенно рассматриваться как рассеяние медленных электронов с массой m^* на атоме водорода, погруженного в среду с диэлектрической постоянной ε_0 [2]).

Если скорость рассеиваемых электронов мала, то рассеяние сферически симметрично, и в квантовомеханических формулах теории рассеяния достаточно учесть только нулевую фазу. Если учитывать эффекты электронного обмена и поляризации рассеивающего атома, то расчет сечения может быть проведен только

---

[1]) Коган В. И., Галицкий В. М. Сборник задач по квантовой механике, М.: Гостехиздат, гл. XI, задача 1.

[2]) Erginsoy C. — Phys. Rev., 1950, v. 79, p. 1013.

численно. При этом для времени релаксации получается результат

$$\frac{1}{\tau} = \frac{20\varepsilon_0\hbar^3}{m^{*2}e^2} n_0, \tag{7.15}$$

где n_0 — концентрация нейтральных атомов примеси.

Отметим, что во время релаксации для рассеяния на нейтральных примесях не зависит от энергии. Как будет показано ниже, отсюда следует, что подвижность носителей тока не зависит от температуры (если, конечно, n_0 не зависит от температуры). Можно показать, что в некоторых случаях рассеяние на нейтральных атомах становится сравнимым с рассеянием на ионах примеси.

4. Рассмотрим вопрос об определении времени релаксации τ в случае, когда одновременно действует несколько механизмов рассеяния.

Для определенности представим себе, что электрон проводимости обладает вероятностью $W_L(\boldsymbol{k}, \boldsymbol{k}')$ в единицу времени перейти из состояния $\boldsymbol{k}$ в состояние $\boldsymbol{k}'$ в результате взаимодействия с акустическими колебаниями и вероятностью $W_I(\boldsymbol{k}, \boldsymbol{k}')$ рассеяться на ионах примеси. Если оба процесса альтернативны, т. е. может произойти либо один из них, либо другой, то полная вероятность рассеяния в единицу времени $W(\boldsymbol{k}, \boldsymbol{k}') = W_L(\boldsymbol{k}, \boldsymbol{k}') + W_I(\boldsymbol{k}, \boldsymbol{k}')$.

Если столкновения упруги, то можно ввести время релаксации по формуле (2.12):

$$\frac{1}{\tau} = -\sum_{\boldsymbol{k}'} W(\boldsymbol{k}, \boldsymbol{k}') \frac{\Delta k_\chi}{k_\chi} =$$

$$= -\sum_{\boldsymbol{k}'} W_L(\boldsymbol{k}, \boldsymbol{k}') \frac{\Delta k_\chi}{k_\chi} - \sum_{\boldsymbol{k}'} W_I(\boldsymbol{k}, \boldsymbol{k}') \frac{\Delta k_\chi}{k_\chi} = \frac{1}{\tau_L} + \frac{1}{\tau_I}. \tag{7.16}$$

Таким образом, если одновременно действует несколько механизмов рассеяния, то для определения обратного значения результирующего (эффективного) времени релаксации складываются обратные же значения времени релаксаций, соответствующие отдельным механизмам рассеяния.

В общем случае

$$\frac{1}{\tau} = \sum_i \frac{1}{\tau_i}, \tag{7.17}$$

где τ_i — время релаксации, определяемое i-м механизмом рассеяния. Во всех случаях, рассмотренных в настоящей главе, время релаксации может быть выражено в виде

$$\tau = \tau_0 \varepsilon^r, \tag{7.18}$$

где показатель степени r имеет различные значения для разных механизмов рассеяния. Например, в случае рассеяния на акустических колебаниях и ионах примеси

$$\tau_L = \frac{\tau_{0L}}{k_0 T} \varepsilon^{-1/2}, \quad \tau_I = \tau_{0I} \varepsilon^{3/2}, \tag{7.18a}$$

т. е. r равно соответственно $-1/2$ и $3/2$.

Согласно (7.16) эффективное время релаксации при действии обоих механизмов равно

$$\tau = \frac{\tau_{0L}\tau_{0I}\varepsilon^{3/2}}{\tau_{0L} + \tau_{0I}(k_0 T)\varepsilon^2} \tag{7.19}$$

и, как мы видим, уже не имеет простой структуры (7.18).

Отметим, что если какой-либо механизм рассеяния перестает действовать, то его время релаксации $\tau \longrightarrow \infty$. Так, если рассеяние на ионах примеси несущественно, то $\tau_{0I} \longrightarrow \infty$ и из (7.19) следует

$$\tau = \frac{\tau_{0L}}{k_0 T} \varepsilon^{-1/2} = \tau_L. \tag{7.19a}$$

ГЛАВА IX

КИНЕТИЧЕСКИЕ ПРОЦЕССЫ (ЯВЛЕНИЯ ПЕРЕНОСА) В ПОЛУПРОВОДНИКАХ

§ 1. Введение

1. Кинетические процессы или явления переноса, осуществляемые электронами и дырками в твердых телах, весьма интересны и важны как с теоретической, так и с практической точек зрения.

Как уже отмечалось выше (гл. VIII, § 1), кинетические процессы, в противоположность статистически равновесным состояниям, зависят от механизма взаимодействия внутри системы. Это делает теорию кинетических процессов более сложной и менее достоверной, чем теорию равновесных систем. Однако изучение кинетических процессов позволяет исследовать механизмы взаимодействий в системе и такие ее структурные свойства, которые не могут быть определены из изучения равновесных систем.

С другой стороны, и практическое применение полупроводников основано на использовании их в качестве проводников электронного и дырочного токов, т. е. на явлениях переноса. В современных установках, использующих полупроводниковые материалы, наряду с электрическими полями широко применяются магнитные поля, градиенты температуры и градиенты концентрации носителей тока; поэтому существует практическая потребность изучения всех разнообразных кинетических процессов: электро- и теплопроводности, термоэлектрических и гальваномагнитных явлений и т. д.

При изучении кинетических процессов из-за сложного характера условий, от которых они зависят, возникает большое число разнообразных случаев.

Во-первых, велико само число даже основных кинетических эффектов. Перечислим только главнейшие из них: электро- и теплопроводность, термоэлектродвижущаяся сила, эффекты Томсона и Пельтье, эффект Холла, изменение сопротивления в магнитном поле, поперечный и продольный эффекты Нернста.

Во-вторых, в ряде случаев приходится иметь дело с несколькими сортами носителей тока (электроны, дырки, тяжелые и легкие дырки).

В-третьих, необходимо учитывать различные механизмы рассеяния носителей тока: колебания атомной решетки, ионизованные и нейтральные примеси, колебания ионной решетки и др.

В-четвертых, необходимо различать случаи, когда электроны проводимости находятся в вырожденном, когда в невырожденном состоянии.

В-пятых, большое значение имеет характер зависимости энергии электрона (дырки) от волнового вектора $\boldsymbol{k}$; поверхности постоянной энергии могут быть: сферами, эллипсоидами, гофрированными поверхностями (гл. IV, § 15).

Если добавить к этому проблему «увлечения» электронов проводимости неравновесным распределением фононов в термоэлектрических и термомагнитных явлениях и проблему «горячих» электронов (сильные электрические поля), то общее число различных случаев, которые могут быть рассмотрены, превышает несколько сот. Рассматривать все эти теоретически возможные случаи не только невозможно, но и нецелесообразно. Развитие физики полупроводников за последние 20 лет показало чрезвычайную сложность и большое разнообразие их внутренней структуры, поэтому заведомо бесполезно пытаться описать все известные в настоящее время и возможные в будущем случаи.

Мы рассмотрим явления переноса в полупроводниках, руководствуясь следующей схемой:

1) невырожденные полупроводники с простой зоной энергии;

2) полупроводники с простой зоной энергии при учете вырождения носителей тока;

3) полупроводники типа германия;

4) полупроводники со сферической непараболической зоной энергии (модель Кейна);

5) эффект «увлечения» носителей тока фононами;

6) квантовая теория гальвано- и термомагнитных явлений.

2. Для того чтобы ток электронов проводимости в полупроводнике равнялся нулю, необходимо, чтобы они находились в состоянии термодинамического равновесия. Для этого требуется, чтобы не только их температура T во всех точках была постоянной, но и был бы постоянен их химический потенциал $\zeta^* = \zeta - e\varphi$, где ζ — так называемый электрохимический потенциал, определяемый концентрацией и температурой электронов в данной точке полупроводника, а $\varphi(x, y, z)$ — электростатический потенциал в этой же точке ($-e\varphi$ — энергия электрона в поле этого потенциала) [1]).

---

[1]) То, что в статистическом равновесии системы $\zeta^* = \text{const}$, вытекает из общих положений статистической физики (см., например, Ансельм А. И., стр. 219). До настоящего времени (например, гл. VI, § 2) мы пользовались электрохимическим потенциалом ζ, называя его просто химическим потенциалом, так мы будем делать и в дальнейшем.

Электронный ток возникает в результате того, что T или $\frac{\zeta}{e}-\varphi=\frac{\zeta^*}{e}$ не постоянны в объеме полупроводника. Очевидно, что в стационарном случае в изотропном полупроводнике в линейном приближении электронный ток состоит из слагаемых, пропорциональных ∇T, и $\nabla\left(\frac{\zeta}{e}-\varphi\right)$, т. е. плотность постоянного тока

$$\boldsymbol{j}=\sigma\nabla\left(\frac{\zeta}{e}-\varphi\right)-\beta\nabla T. \tag{1.1}$$

В однородном полупроводнике в изотермических условиях ζ постоянно по его объему, поэтому $\nabla\left(\frac{\zeta}{e}-\varphi\right)=-\nabla\varphi=\boldsymbol{E}$, т. е. совпадает с напряженностью электростатического поля $\boldsymbol{E}$. Если неоднородный полупроводник находится в состоянии статистического равновесия ($\nabla T=0$, $\boldsymbol{j}=0$), то в нем возникают электрические поля с напряженностью $\boldsymbol{E}=-\nabla(\zeta/e)$. Соотношение (1.1) носит локальный характер, т. е. все величины, входящие в (1.1), вообще говоря, зависят от координат выбранной точки $\boldsymbol{r}$.

Наряду с плотностью электрического тока $\boldsymbol{j}$ рассмотрим плотность потока «кинетической» или тепловой энергии $\boldsymbol{q}=\boldsymbol{w}-\left(\varphi-\frac{\zeta}{e}\right)\boldsymbol{j}$, где из плотности потока полной энергии $\boldsymbol{w}$ вычтена часть, связанная с «потенциальной» энергией, переносимой потоком электронов. В линейном приближении

$$\boldsymbol{q}=\boldsymbol{w}-\left(\varphi-\frac{\zeta}{e}\right)\boldsymbol{j}=\gamma\nabla\left(\frac{\zeta}{e}-\varphi\right)-\varkappa\nabla T. \tag{1.2}$$

В случае анизотропного полупроводника скалярные кинетические коэффициенты σ, β, γ, $\varkappa$ превращаются в тензоры второго ранга и (1.1) и (1.2) приобретают вид

$$j_i=\sum_k\sigma_{ik}\nabla_k\left(\frac{\zeta}{e}-\varphi\right)-\sum_k\beta_{ik}\nabla_k T, \tag{1.1a}$$

$$q_i=\sum_k\gamma_{ik}\nabla_k\left(\frac{\zeta}{e}-\varphi\right)-\sum_k\varkappa_{ik}\nabla_k T. \tag{1.2a}$$

Здесь индексы i, $k=1$, 2, 3 или i, $k=x$, y, z. Коэффициенты, входящие в (1.1a) и (1.2a), удовлетворяют соотношениям симметрии или принципу Онсагера[1]):

$$\sigma_{ik}=\sigma_{ki}, \tag{1.3}$$

$$\varkappa_{ik}=\varkappa_{ki}, \tag{1.4}$$

$$\gamma_{ik}=T\beta_{ki}. \tag{1.5}$$

Если полупроводник помещен в магнитное поле, то кинетические коэффициенты зависят от напряженности магнитного поля $\boldsymbol{H}$;

[1]) Ансельм А. И., гл. X, § 6.

в этом случае соотношения (1.3)—(1.5) обобщаются следующим образом:

$$\sigma_{ik}(\boldsymbol{H}) = \sigma_{ki}(-\boldsymbol{H}), \tag{1.3a}$$

$$\varkappa_{ik}(\boldsymbol{H}) = \varkappa_{ki}(-\boldsymbol{H}), \tag{1.4a}$$

$$\gamma_{ik}(\boldsymbol{H}) = T\beta_{ki}(-\boldsymbol{H}). \tag{1.5a}$$

Особенно полезным является соотношение (1.5а), позволяющее заменить вычисление электрического тока при наличии градиента температуры, определением потока энергии в электрическом поле.

Если носители тока движутся по законам классической механики, то кинетические коэффициенты могут быть в принципе определены из кинетического уравнения. Этому и будет в основном посвящена настоящая глава.

Отметим, что и в случае квантовой теории явлений переноса имеют место соотношения Онсагера (1.3а)—(1.5а).

§ 2. Определение неравновесной функции для электронов проводимости в случае сферически симметричной зоны

1. В случае сферически симметричной зоны энергия электрона

$$\varepsilon = \varepsilon(k) \tag{2.1}$$

зависит от абсолютной величины волнового вектора $k = |\boldsymbol{k}|$, т. е. поверхности постоянной энергии сферы.

В этом случае скорость электрона

$$\boldsymbol{v} = \frac{1}{\hbar}\frac{\partial\varepsilon(k)}{\partial\boldsymbol{k}} = \frac{1}{\hbar}\frac{\partial\varepsilon(k)}{\partial k}\frac{\boldsymbol{k}}{k} = u(k)\,\boldsymbol{k}, \tag{2.2}$$

где

$$u(k) = \frac{1}{\hbar k}\frac{\partial\varepsilon(k)}{\partial k}. \tag{2.2a}$$

Таким образом, в этом случае $\boldsymbol{v} \parallel \boldsymbol{k}$.

Положим в соответствии с (VIII.2.9а) неравновесную функцию распределения

$$f(\boldsymbol{k}) = f_0(\varepsilon) + f_1(\boldsymbol{k}), \quad \text{с } f_1(\boldsymbol{k}) = -\frac{\partial f_0}{\partial\varepsilon}\chi(\varepsilon)\,\boldsymbol{k}, \tag{2.3}$$

где равновесная функция распределения

$$f_0(\varepsilon) = \frac{1}{\exp\dfrac{\varepsilon-\zeta}{k_0T}+1}. \tag{2.3a}$$

Структура члена $f_1(\boldsymbol{k})$ в (2.3) будет обоснована ниже.

Мы определим добавку $f_1(\boldsymbol{k})$ из кинетического уравнения (VIII.2.6) и (VIII.2.11а):

$$\boldsymbol{v}\nabla_r f - \frac{e}{\hbar}\left(\boldsymbol{E} + \frac{1}{c}[\boldsymbol{v}\boldsymbol{H}]\right)\nabla_k f = -\frac{f_1(\boldsymbol{k})}{\tau(\varepsilon)}, \tag{2.4}$$

т. е. в приближении времени релаксации $\tau(\varepsilon)$.

Если мы интересуемся токами, пропорциональными $\boldsymbol{E}$ и ∇T (линейная область явлений переноса), то поправка $f_1(\boldsymbol{k})$ тоже должна определяться в линейном приближении по $\boldsymbol{E}$ и ∇T, а для этого в левую часть уравнения (2.4) надо подставлять $f \approx f_0$.

Таким образом,

$$\nabla_r f \approx \nabla_r f_0 = \frac{\partial f_0}{\partial \varepsilon}\left\{\frac{\zeta-\varepsilon}{T}\nabla T - \nabla\zeta\right\}, \tag{2.5}$$

где мы считаем, что имеет место локальное термодинамическое равновесие, т. е. $T = T(\boldsymbol{r})$ и $\zeta = \zeta(\boldsymbol{r})$.

С другой стороны,

$$\nabla_k f \approx \nabla_k f_0 = \frac{\partial f_0}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial \boldsymbol{k}} = \frac{\partial f_0}{\partial \varepsilon}\hbar\boldsymbol{v}. \tag{2.6}$$

Мы видим, что в приближении $f \approx f_0$ слагаемое, содержащее магнитное поле в левой части (2.4), выпадает, так как $[\boldsymbol{vH}]\boldsymbol{v} \equiv 0$. Таким образом, для того чтобы учесть магнитное поле в кинетическом уравнении (2.4), надо вычислить $\nabla_k f$ в следующем приближении:

$$\nabla_k f = \nabla_k\left[f_0 - \frac{\partial f_0}{\partial \varepsilon}\boldsymbol{\chi}(\varepsilon)\boldsymbol{k}\right] = \frac{\partial f_0}{\partial \varepsilon}\hbar\boldsymbol{v} - (\boldsymbol{\chi k})\nabla_k\left(\frac{\partial f_0}{\partial \varepsilon}\right) - \frac{\partial f_0}{\partial \varepsilon}\nabla_k(\boldsymbol{\chi k}). \tag{2.6а}$$

Так как

$$\nabla_k\left(\frac{\partial f_0}{\partial \varepsilon}\right) = \left(\frac{\partial^2 f_0}{\partial \varepsilon^2}\right)\nabla_k\varepsilon = \left(\frac{\partial^2 f_0}{\partial \varepsilon^2}\right)\hbar\boldsymbol{v}, \tag{2.6б}$$

то в (2.6а) не только первое, но и второе слагаемое дает нуль при умножении на $[\boldsymbol{vH}]$.

Для определения третьего слагаемого в правой части (2.6а) вычислим составляющую $\nabla_k(\boldsymbol{\chi k})$ по оси k_x

$$\frac{\partial}{\partial k_x}(\boldsymbol{\chi k}) = \frac{\partial}{\partial k_x}(k_x\chi_x + k_y\chi_y + k_z\chi_z) = \chi_x + \left(k_x\frac{\partial \chi_x}{\partial k_x} + k_y\frac{\partial \chi_y}{\partial k_x} + k_z\frac{\partial \chi_z}{\partial k_x}\right) =$$

$$= \chi_x + \left(k_x\frac{\partial \chi_x}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial k_x} + k_y\frac{\partial \chi_y}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial k_x} + k_z\frac{\partial \chi_z}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial k_x}\right) =$$

$$= \chi_x + \left(k_x\frac{\partial \chi_x}{\partial \varepsilon} + k_y\frac{\partial \chi_y}{\partial \varepsilon} + k_z\frac{\partial \chi_z}{\partial \varepsilon}\right)\hbar v_x.$$

Таким образом

$$\nabla_k(\boldsymbol{\chi k}) = \boldsymbol{\chi} + \left(k_x\frac{\partial \chi_x}{\partial \varepsilon} + k_y\frac{\partial \chi_y}{\partial \varepsilon} + k_z\frac{\partial \chi_z}{\partial \varepsilon}\right)\hbar\boldsymbol{v}. \tag{2.6в}$$

Подставляя (2.5), (2.6а)—(2.6в) в (2.4), получим

$$\frac{\partial f_0}{\partial \varepsilon}\left\{\frac{\zeta-\varepsilon}{T}\nabla T - \nabla\zeta - e\boldsymbol{E} - \frac{e}{\hbar c}[\boldsymbol{H\chi}]\right\}\boldsymbol{v} = \frac{1}{\tau(\varepsilon)}\left(\frac{\partial f_0}{\partial \varepsilon}\right)\frac{1}{u(k)}\boldsymbol{\chi v}. \tag{2.7}$$

В левой части мы использовали циклическое свойство скалярно-векторного произведения[1]): $[\boldsymbol{vH}]\boldsymbol{\chi} = [\boldsymbol{H\chi}]\boldsymbol{v}$; в правой — заменили

[1]) Смирнов В. И., т. II, § 105.

$f_1(\boldsymbol{k})$ на (2.3) и положили $\boldsymbol{k}=\boldsymbol{v}/u(k)$. Выражение (2.7) оправдывает форму (2.3) для $f_1(\boldsymbol{k})$. Так как $\boldsymbol{v}$ произвольно, то из (2.7) следует для электронов проводимости

$$\chi_n(\varepsilon)=-\tau_n(\varepsilon)\,u_n(k)\left\{\frac{\varepsilon-\zeta}{T}\nabla T+\nabla(\zeta-e\varphi)-\frac{e}{\hbar c}[\boldsymbol{H}\chi_n]\right\}. \quad (2.8)$$

Здесь $\varphi(\boldsymbol{r})$ — электростатический потенциал приложенного электрического поля $\boldsymbol{E}=-\nabla\varphi$; индекс n здесь и далее отмечает, что соответствующие величины относятся к электронам (negative).

В (гл. VI, § 2, п. 4) мы видели, что статистическое поведение дырок вполне эквивалентно поведению электронов, если приписать им энергию ε' и химический потенциал ζ', равные

$$\varepsilon'=\frac{\hbar^2k'^2}{2m_p}, \quad (2.9)$$

$$\zeta'=-\varepsilon_G-\zeta, \quad (2.9a)$$

где ε_G — ширина запрещенной зоны. Тогда для дырок получим вместо (2.8)

$$\chi_p(\varepsilon')=-\tau_p(\varepsilon')\,u_p(k')\left\{\frac{\varepsilon'+\varepsilon_G+\zeta}{T}\nabla T-\nabla(\zeta-e\varphi)+\frac{e}{\hbar c}[\boldsymbol{H}\chi_p]\right\}, \quad (2.10)$$

причем по сравнению с (2.8) для положительных дырок изменен знак заряда. Индекс p указывает, что соответствующие величины характеризуют положительные дырки (positive).

Для того чтобы (2.8) и (2.10) определяли неравновесные функции распределения для электронов и дырок, необходимо, чтобы времена релаксации τ_n и τ_p были много больше среднего времени жизни электрона (дырки) по отношению к процессам рекомбинации.

Если магнитное поле $\boldsymbol{H}=0$, то (2.8) и (2.10) непосредственно определяют χ_n и χ_p; при $\boldsymbol{H}\neq 0$ соответствующие уравнения должны быть решены относительно χ_n и χ_p. Можно показать (Приложение 23), что в этом случае

$$\chi_n=-\tau_nu_n\frac{\left\{\frac{\varepsilon-\zeta}{T}\nabla T+\nabla(\zeta-e\varphi)\right\}+\frac{e\tau_nu_n}{\hbar c}\left[\boldsymbol{H}\left\{\frac{\varepsilon-\zeta}{T}\nabla T+\nabla(\zeta-e\varphi)\right\}\right]}{1+\left(\frac{e\tau_nu_n}{\hbar c}\boldsymbol{H}\right)^2}+$$

$$+\frac{\left(\frac{e\tau_nu_n}{\hbar c}\right)^2\left(\left\{\frac{\varepsilon-\zeta}{T}\nabla T+\nabla(\zeta-e\varphi)\right\}\boldsymbol{H}\right)\boldsymbol{H}}{1+\left(\frac{e\tau_nu_n}{\hbar c}\boldsymbol{H}\right)^2} \quad (2.11)$$

и

$$\boldsymbol{\chi}_p = -\tau_p u_p \times$$
$$\times \frac{\left\{\frac{\varepsilon'+\varepsilon_G+\zeta}{T}\nabla T - \nabla(\zeta - e\varphi)\right\} - \frac{e\tau_p u_p}{\hbar c}\left[\boldsymbol{H}\left\{\frac{\varepsilon'+\varepsilon_G+\zeta}{T}\nabla T - \nabla(\zeta-e\varphi)\right\}\right]}{1+\left(\frac{e\tau_p u_p}{\hbar c}\boldsymbol{H}\right)^2} +$$
$$+ \frac{\left(\frac{e\tau_p u_p}{\hbar c}\right)^2 \left(\left\{\frac{\varepsilon'+\varepsilon_G+\zeta}{T}\nabla T - \nabla(\zeta-e\varphi)\right\}\boldsymbol{H}\right)\boldsymbol{H}}{1+\left(\frac{e\tau_p u_p}{\hbar c}\boldsymbol{H}\right)^2}. \qquad (2.12)$$

Эти сложные выражения, как мы увидим ниже, значительно упрощаются в частных случаях, когда магнитное поле $\boldsymbol{H}=0$ или оно слабо (точное определение слабого магнитного поля дано ниже), когда градиент температуры $\nabla T = 0$, и т. п.

В случае простой зоны энергия электрона

$$\varepsilon = \hbar^2 k^2 / 2m_n, \qquad (2.13)$$

поэтому, как это следует из (2.2а),

$$u_n = \frac{1}{\hbar k}\frac{\partial \varepsilon}{\partial k} = \frac{\hbar}{m_n} \qquad (2.13а)$$

и аналогично для дырок u_p.

2. Вычислим электрический ток, создаваемый неравновесно распределенными электронами. Из (VII.2.14) следует, что плотность тока

$$\boldsymbol{j}_n = -\frac{e}{4\pi^3}\int \boldsymbol{v} f_1(\boldsymbol{k})\, d^3k = -\frac{e}{4\pi^3}\int\left[-\frac{\partial f_0}{\partial \varepsilon}u_n(k)\right](\boldsymbol{\chi}_n \boldsymbol{k})\,\boldsymbol{k}\, d^3k, \quad (2.14)$$

где мы воспользовались (2.2) и (2.3); здесь $\boldsymbol{\chi}_n$ равно (2.11). Интеграл, стоящий в (2.14) в полярных координатах (полярная ось $z \| \boldsymbol{\chi}_n$), равен

$$\int\left[-\frac{\partial f_0}{\partial \varepsilon}u_n(k)\right]\chi_n k\cos\vartheta\,\{\boldsymbol{i}_0 k\sin\vartheta\cos\varphi + \boldsymbol{j}_0 k\sin\vartheta\sin\varphi +$$
$$+ \boldsymbol{k}_0 k\cos\vartheta\}\, k^2\, dk\sin\vartheta\, d\vartheta\, d\varphi$$

Здесь $\boldsymbol{i}_0$, $\boldsymbol{j}_0$, $\boldsymbol{k}_0$ — единичные векторы, направленные вдоль осей x, y, z. При интегрировании по φ от 0 до 2π первое и второе слагаемые в фигурных скобках дают нуль; поэтому интегрирование по φ и ϑ дает

$$\int_0^{2\pi} d\varphi \int_0^{\pi}\cos^2\vartheta\sin\vartheta\, d\vartheta = \frac{4\pi}{3}.$$

Таким образом, получим для (2.14)

$$\boldsymbol{j}=-\frac{e}{3\pi^2}\int_0^\infty\left[-\frac{\partial f_0}{\partial\varepsilon}u_n(k)\right]\chi_n(k)\,k^4\,dk. \tag{2.15}$$

Для тока, создаваемого дырками, надо в этом выражении заменить $-e$ на e, f_0—на равновесную функцию распределения дырок и χ_n на χ_p (2.12).

Для электронов в случае простой зоны энергия ε и $u_n(k)$ равны (2.13) и (2.13а). Положим в этом случае

$$\chi_n=-\frac{\hbar e}{m_n}\chi_n^*, \tag{2.16}$$

где χ_n равно (2.11); тогда

$$\chi_n^*=\frac{\tau_n\boldsymbol{P}_n+\gamma_n\tau_n^2\,[\boldsymbol{H}\boldsymbol{P}_n]+\gamma_n^2\tau_n^3\,(\boldsymbol{H}\boldsymbol{P}_n)\,\boldsymbol{H}}{1+(\gamma_n\tau_nH)^2} \tag{2.17}$$

с

$$\boldsymbol{P}_n=\frac{\varepsilon-\zeta}{eT}\nabla T+\nabla\left(\frac{\zeta}{e}-\varphi\right) \tag{2.17а}$$

и

$$\gamma_n=e/m_nc. \tag{2.17б}$$

Вычислим электрический ток, создаваемый «полем» χ_n, равным (2.16). Для невырожденного полупроводника, как это следует из (VI.2.11а),

$$\frac{\partial f_0}{\partial\varepsilon}=\frac{4\pi^3\hbar^3n}{(2\pi m_n)^{3/2}\,(k_0T)^{5/2}}\,e^{-\varepsilon/k_0T}. \tag{2.18}$$

Подставляя это значение в интеграл (2.15), переходя в нем посредством (2.13) от переменной интегрирования k к $x=\varepsilon/k_0T$ и используя (2.16), получим

$$\boldsymbol{j}_n=\frac{ne^2}{m_n}\frac{4}{3\sqrt{\pi}}\int_0^\infty\chi_n^*e^{-x}x^{3/2}\,dx. \tag{2.19}$$

Удобно ввести следующий символ усреднения:

$$\langle\chi_n^*\rangle=\frac{4}{3\sqrt{\pi}}\int_0^\infty\chi_n^*e^{-x}x^{3/2}dx. \tag{2.20}$$

Если $\boldsymbol{g}=\text{const}$, то

$$\langle\boldsymbol{g}\rangle=\boldsymbol{g}\frac{4}{3\sqrt{\pi}}\int_0^\infty e^{-x}x^{3/2}\,dx=\boldsymbol{g}\frac{4}{3\sqrt{\pi}}\Gamma\left(\frac{5}{2}\right)=\boldsymbol{g}, \tag{2.20а}$$

как это следует из (П.7.12).

Таким образом, плотность тока

$$\boldsymbol{j}_n=\frac{ne^2}{m_n}\langle\chi_n^*\rangle. \tag{2.21}$$

Электрический ток, обусловленный дырками, вычисляется аналогично. Если ввести для дырок вектор χ_p^*, связанный с χ_p (2.12) соотношением

$$\chi_p = \frac{\hbar e}{m_p}\chi_p^*, \tag{2.22}$$

то

$$\chi_p^* = \frac{\tau_p \boldsymbol{P}_p - \gamma_p \tau_p^2 [\boldsymbol{H}\boldsymbol{P}_p] + \gamma_p^2 \tau_p^3 (\boldsymbol{H}\boldsymbol{P}_p)\boldsymbol{H}}{1 + (\gamma_p \tau_p H)^2}, \tag{2.23}$$

где

$$\boldsymbol{P}_p = -\frac{\varepsilon' + \varepsilon_G + \zeta}{eT}\nabla T + \nabla\left(\frac{\zeta}{e} - \varphi\right) \tag{2.23а}$$

и

$$\gamma_p = e/m_p c. \tag{2.23б}$$

Так как для дырок интеграл (2.15) имеет противоположный знак, но связь между χ_p и χ_p^* (2.22) не содержит знак минус, подобно (2.16), то плотность тока

$$\boldsymbol{j}_p = \frac{pe^2}{m_p}\langle\chi_p^*\rangle, \tag{2.24}$$

аналогично (2.21). Общий ток, обусловленный электронами и дырками, равен

$$\boldsymbol{j} = \boldsymbol{j}_n + \boldsymbol{j}_p = \frac{ne^2}{m_n}\langle\chi_n^*\rangle + \frac{pe^2}{m_p}\langle\chi_p^*\rangle, \tag{2.25}$$

где χ_n^* и χ_p^* равны (2.17) и (2.23).

§ 3. Электропроводность невырожденных полупроводников с простой зонной структурой

Рассмотрим электрический ток в однородном электронном полупроводнике в отсутствие градиента температуры и магнитного поля ($\nabla T = 0$, $\nabla\zeta = 0$, $\boldsymbol{H} = 0$); в этом случае

$$\chi^* = \tau\boldsymbol{P} = \tau(-\nabla\varphi) = \tau\boldsymbol{E}, \tag{3.1}$$

где $\boldsymbol{E}$ — напряженность электрического поля [1]).

Из (2.21) и (3.1) следует, что

$$\boldsymbol{j} = \frac{ne^2}{m}\langle\tau\rangle\boldsymbol{E} = \sigma\boldsymbol{E}, \tag{3.2}$$

где σ — удельная электропроводность. Отсюда и из (VI.7.7а) следует, что подвижность электронов

$$\mu = \frac{\sigma}{en} = \frac{e}{m}\langle\tau\rangle. \tag{3.3}$$

---

[1]) Для упрощения записи мы у соответствующих величин в выражениях (3.1)—(3.6) опускаем индекс n.

Если время релаксации τ от энергии ε не зависит, то

$$\mu = e\tau/m, \tag{3.3а}$$

как это следует из (2.20а). Если ввести среднюю дрейфовую скорость носителей тока $\boldsymbol{v}_{др}$, положив по определению

$$\boldsymbol{j} = \sigma \boldsymbol{E} = en\boldsymbol{v}_{др},$$

то из сравнения с (3.3) следует

$$\mu = v_{др}/E,$$

т. е. подвижность численно равна дрейфовой скорости (но не имеет размерности скорости!) при напряженности электрического поля E, равной единице.

В предыдущей главе мы рассмотрели ряд механизмов рассеяния электронов в кристаллах, для которых время релаксации имеет вид

$$\tau = a\varepsilon^r = a(k_0T)^r x^r = \tau_0 x^r, \tag{3.4}$$

где $x = \varepsilon/k_0T$ и $\tau_0 = a(k_0T)^r$ от x не зависит. Из (3.4), (2.20) и Приложения 7, п. 2 следует, что

$$\langle\tau\rangle = \tau_0 \frac{4}{3\sqrt{\pi}} \Gamma\left(\frac{5}{2}+r\right). \tag{3.4а}$$

Из (VIII.4.11), (VIII.7.14), (VIII.6.7), (VIII.6.13) следует, что для рассеяния электронов на:

а) акустических колебаниях $\tau = a_A (k_0T)^{-3/2} x^{-1/2}$; (3.5а)

б) ионах примеси $\tau = a_I (k_0T)^{3/2} x^{3/2}$; (3.5б)

в) оптических колебаниях в ионных кристаллах выше температуры Дебая $\tau = a_P (k_0T)^{-1/2} x^{1/2}$; (3.5в)

г) и ниже температуры Дебая $\tau = a_{0P} \exp\left(\frac{\hbar\omega_0}{k_0T}\right)$. (3.5г)

Для области эффективного пьезоэлектрического рассеяния зависимость времени релаксации τ от T и $x = \varepsilon/k_0T$ такая же, как в случае в). Значения a_A, a_I, a_P, и a_{0P} могут быть определены из сравнения выражений (3.5а)—(3.5г) с соответствующими выражениями гл. VIII.

Подставим (3.5а)—(3.5г) в (3.3); используя Приложение 7, п. 2, получим:

а) $$\mu = \frac{4}{3\sqrt{\pi}} \frac{e}{m} \frac{a_A}{(k_0T)^{3/2}} \sim T^{-3/2}; \tag{3.6а}$$

б) $$\mu = \frac{8}{\sqrt{\pi}} \frac{e}{m} a_I (k_0T)^{3/2} \sim T^{3/2}; \tag{3.6б}$$

в) $$\mu = \frac{8}{3\sqrt{\pi}} \frac{e}{m} \frac{a_P}{(k_0T)^{1/2}} \sim T^{-1/2}; \tag{3.6в}$$

г) $$\mu = \frac{e}{m} a_{0P} \exp\left(\frac{\hbar\omega_0}{k_0T}\right) \sim \exp\left(\frac{\hbar\omega_0}{k_0T}\right). \tag{3.6г}$$

Если одновременно действуют несколько механизмов рассеяния, то время релаксации, которое нужно подставить в (3.2), определяется по формуле (VIII.7.17).

Сравнивая (VIII.4.11) с (3.6а), получим для подвижности, обусловленной рассеянием на акустических колебаниях,

$$\mu = 3\sqrt{\frac{\pi}{2}}\frac{e\rho v_0^2\hbar^4}{C^2 m^{*5/2}(k_0T)^{3/2}}, \tag{3.7}$$

где плотность кристалла $\rho = M/\Omega_0$. Конечно ρ, v_0, C и m^* зависят от температуры, но зависимость эта слабая, поэтому в основном подвижность $\mu \sim T^{-3/2}$.

Электрический ток, обусловленный дырками, вычисляется по формуле (2.24). Полный ток

$$\boldsymbol{j} = e(n\mu_n + p\mu_p)\boldsymbol{E}, \tag{3.8}$$

где n и p — концентрации электронов и дырок, а подвижность дырок μ_p, обусловленная, например, взаимодействием с акустическими колебаниями, определяется выражением (3.7), в котором константа взаимодействия C и эффективная масса m^* должны быть взяты для дырок.

Зависимость, близкая к (3.7), наблюдается у n-Ge: $\mu_n \sim T^{-1,66}$ ($100° < T \leqslant 280°$); однако для p-Ge $\mu_p \sim T^{-2,33}$ (в том же интервале температур). Для электронного и дырочного кремния: $\mu_n \sim T^{-2,6}$ ($300° < T \lesssim 400°$) и $\mu_p \sim T^{-2,3}$ ($150° < T < 400°$). Такие отступления от теории, которая для полупроводников с ковалентными связями приводит к закону $\mu \sim T^{-3/2}$, объясняются различными причинами: дополнительным рассеянием на оптических колебаниях решетки, междолинными переходами электронов при их рассеянии и непараболичностью зоны, т. е. зависимостью эффективной массы электрона от его энергии.

При понижении температуры доминирующим механизмом рассеяния становится рассеяние на ионах примеси, которое приводит к зависимости (3.6б): $\mu \sim T^{3/2}$ этот переход наблюдается на опыте, например, у n-Ge.

В некоторых случаях существенны одновременно несколько различных механизмов рассеяния. Например, подвижность, обусловленная взаимодействием с акустическими колебаниями, убывает при возрастании температуры как $T^{-3/2}$, а подвижность, связанная с рассеянием на ионах примеси, возрастает как $T^{3/2}$. Поэтому при высоких и низких температурах можно соответственно пренебречь рассеянием на ионах или взаимодействием с акустическими фононами, однако в промежуточной области температур будут существенны оба механизма рассеяния. Выводы настоящего параграфа справедливы при любой зависимости времени релаксации τ от энергии, поэтому все выведенные выше

формулы остаются в силе, если вместо τ подставлять его эффективное значение, определяемое выражением (VIII.7.17).

Время релаксации, обусловленное одновременным взаимодействием с акустическими колебаниями и ионами примеси, согласно (7.19) равно

$$\tau = \frac{3\sqrt{\pi}}{4}\frac{m^*}{e}\mu_A \frac{x^{3/2}}{x^2+\beta^2}, \tag{3.9}$$

где по-прежнему $x = \varepsilon/k_0T$, $\beta = \sqrt{6\mu_A/\mu_I}$, а μ_A и μ_I — подвижности (3.6а) и (3.6б), которые мы ввели вместо a_A и a_I. Заметим, что $\beta^2 \sim T^{-3}$.

Из (3.3) и из (3.9) следует, что подвижность

$$\mu = \frac{3\sqrt{\pi}}{4}\mu_A \left\langle \frac{x^{3/2}}{x^2+\beta^2} \right\rangle, \tag{3.10}$$

т. е. определяется интегралом

$$\int_0^\infty \frac{x^3 e^{-x}\,dx}{x^2+\beta^2} = J(\beta), \tag{3.11}$$

зависящим от параметра β.

Можно показать [1]), что

$$J(\beta) = 1 + \beta^2 (\mathrm{Si}\,\beta \sin\beta + \mathrm{Ci}\,\beta \cos\beta), \tag{3.12}$$

где

$$\mathrm{Si}\,\beta = -\int_\beta^\infty \frac{\sin x}{x}\,dx, \quad \mathrm{Ci}\,\beta = -\int_\beta^\infty \frac{\cos x}{x}\,dx \tag{3.12а}$$

— интегральный синус и интегральный косинус [2]).

В результате подвижность μ (3.10) может быть представлена как функция температуры во всем интервале, начиная от высоких и вплоть до низких температур. В предельных случаях $\mu_I \to \infty$ и $\mu_A \to \infty$, подвижность μ, согласно (3.10), соответственно равна $\mu = \mu_A$ и $\mu = \mu_I$.

Аналогично может быть рассмотрено действие нескольких механизмов рассеяния и на другие кинетические коэффициенты: дифференциальную термоэдс, гальвано- и термомагнитные коэффициенты, исследованные ниже в §§ 4, 5 и 6. Встречающиеся при этом интегралы приведены в соответствующей литературе [3]).

---

[1]) Ансельм А. И., Клячкин В. И. — ЖЭТФ, 1952, т. 22, с. 297.

[2]) Таблицы имеются, например, в книге Янке Е, Эмде Ф. Таблицы функций с формулами и кривыми, — М.: Физматгиз, 1959.

[3]) Dingle R. B., Doreen Arndt, Roy S. K. — Applied Scientific Research. — Hague, 1956, Sec. B, vol 6, p. 155.

§ 4. Термоэлектрические явления в невырожденных полупроводниках с простой зонной структурой

1. При наличии градиента температуры ∇T в полупроводнике в нем возникает ряд новых явлений, получивших название *термоэлектрических*. Так как средняя энергия, а часто и число носителей тока возрастают при повышении температуры, то градиент температуры вызывает в направлении $-\nabla T$ поток свободных зарядов. В разомкнутой цепи в стационарном состоянии плотность тока во всех точках равна нулю, что означает, что образовавшееся электрическое поле $\boldsymbol{E}$ вызывает в каждой точке полупроводника ток, компенсирующий поток носителей, пропорциональный ∇T. Возникающая при этом в цепи электродвижущая сила называется *термоэлектродвижущей силой* (*термоэдс*). Так как и электроны и дырки диффундируют в полупроводнике от горячего конца к холодному, то термоэдс у собственных полупроводников, вообще говоря, меньше, чем у примесных. В металлах концентрация электронов, а практически и их энергия не зависят от температуры; поэтому термоэдс у металлов мала по сравнению с полупроводниками.

Нельзя вычислять термоэдс между точками (1) и (2) как линейный интеграл вида

$$\int_{(1)}^{(2)} E_s dS = -\int_{(1)}^{(2)} \nabla_s \varphi \, dS = \varphi_1 - \varphi_2, \tag{4.1}$$

где φ — электростатический потенциал, так как на границе полупроводника и металлического контакта существует контактная разность потенциалов, зависящая от температуры. Таким образом, показания прибора отметят и разность между контактными потенциалами в точках (1) и (2), находящихся при разных температурах. Для того чтобы учесть этот эффект, надо вычислять линейный интеграл не от $-\nabla\varphi$, как в (4.1), а от $\nabla\left(\frac{\zeta}{e}-\varphi\right)$, где ζ — химический потенциал свободных зарядов. В самом деле, полный химический потенциал $(\zeta - e\varphi)$, учитывающий внешнее электрическое поле, не меняется при прохождении контакта, если считать, что температуры полупроводника и металла на контакте одинаковы и что они вблизи границы находятся в статистическом равновесии. На рис. IX.1 кривая *а* изображает ход электростатического потенциала φ в цепи, состоящей из полупроводника AB и металлических подводов OA и BC, вдоль которых падение потенциала практически равно нулю. Так как контакты A и B находятся при разных температурах T_1 и T_2, то контактные скачки потенциала в A и B различны; поэтому

разность потенциалов $V_1 - V_2$, измеряемая прибором, не будет равна разности $\varphi_1 - \varphi_2$, вычисляемой по (4.1).

Кривая *б* изображает ход величины $\varphi - \zeta/e$, которая непрерывна в точках A и B. Так как, с другой стороны, ζ на участках OA и BC одинаково, то разность величины $\varphi - \zeta/e$ в точках A и B равна $V_1 - V_2$, измеряемой прибором.

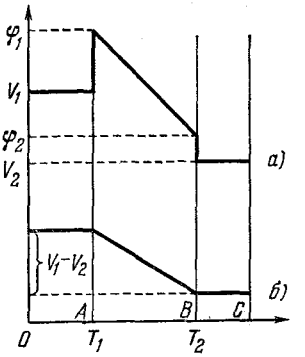


Рис. IX.1.

Наряду с термоэдс существенный интерес представляют еще два термоэлектрических эффекта.

При наличии градиента температуры в проводнике в нем при прохождении тока, наряду с теплом Джоуля, пропорциональным квадрату силы тока j^2, выделяется (поглощается) тепло Томсона, пропорциональное $(j\nabla T)$.

Наконец, при прохождении тока по неоднородному проводнику даже в отсутствие градиента температуры происходит выделение или поглощение тепла (эффект Пельтье). Это явление имеет место, например, при прохождении тока сквозь контакт полупроводника с металлом.

2. В отсутствие магнитного поля ($\boldsymbol{H} = 0$) и наличии градиента температуры ($\nabla T \neq 0$, $\nabla\zeta \neq 0$) из (2.17) следует

$$\chi_n^* = \tau_n P_n = \tau_n \frac{\varepsilon - \zeta}{eT} \nabla T + \tau_n \nabla \left(\frac{\zeta}{e} - \varphi \right) =$$

$$= \frac{k_0}{e} \left[x - \frac{\zeta}{k_0 T} \right] \tau_n \nabla T + \tau_n \nabla \left(\frac{\zeta}{e} - \varphi \right), \qquad (4.2)$$

где $x = \varepsilon/k_0 T$. Из (2.21) следует, что электрический ток, создаваемый неравновесным распределением электронов (4.2), равен

$$\boldsymbol{j}_n = \frac{ne^2}{m_n} \langle \chi_n^* \rangle = \frac{ne^2}{m_n} \left\{ \frac{k_0}{e} \left[\langle \tau_n x \rangle - \frac{\zeta}{k_0 T} \langle \tau_n \rangle \right] \nabla T + \right.$$

$$\left. + \nabla \left(\frac{\zeta}{e} - \varphi \right) \langle \tau_n \rangle \right\} = ne\mu_n \left\{ \frac{k_0}{e} \left(g_n - \frac{\zeta}{k_0 T} \right) \nabla T + \nabla \left(\frac{\zeta}{e} - \varphi \right) \right\}, \qquad (4.3)$$

где подвижность электронов μ_n определяется выражением (3.3), а

$$g_n = \frac{\langle \tau_n x \rangle}{\langle \tau_n \rangle}. \qquad (4.3а)$$

Если время релаксации $\tau_n \sim \varepsilon^r$, то

$$g_n = \frac{\Gamma(r + 7/2)}{\Gamma(r + 5/2)} = r + \frac{5}{2}. \qquad (4.3б)$$

В случае четырех механизмов рассеяния (3.5) величина g_n равна:

а) $g_n = 2$; б) $g_n = 4$; в) $g_n = 3$; г) $g_n = 2{,}5$. (4.4)

Дифференциальная термоэдс α определяется как отношение $\frac{\left|\nabla\left(\frac{\zeta}{e}-\varphi\right)\right|}{|\nabla T|}$ для разомкнутой цепи; из (4.3) при $j_n=0$ получим

$$\alpha=\frac{\left|\nabla\left(\frac{\zeta}{e}-\varphi\right)\right|}{|\nabla T|}=\frac{k_0}{e}\left(g_n-\frac{\zeta}{k_0T}\right)=\frac{k_0}{e}\left[g_n+\lg\frac{2\,(2\pi m_n k_0T)^{3/2}}{nh^3}\right], \tag{4.5}$$

где ζ взято из (VI.2.15б) и $h=2\pi\hbar$. Выражение (4.5) для атомных полупроводников (с $g_n=2$) было впервые получено Н. Л. Писаренко (1940).

Для дырок, как следует из (2.23),

$$\chi_p^*=-\tau_p\frac{\varepsilon'+\varepsilon_G+\zeta}{eT}\nabla T+\tau_p\nabla\left(\frac{\zeta}{e}-\varphi\right). \tag{4.6}$$

Подставляя это выражение в (2.24), получим для электрического тока, обусловленного дырками,

$$j_p=-pe\mu_p\left\{\frac{k_0}{e}\left(g_p+\frac{\varepsilon_G+\zeta}{k_0T}\right)\nabla T-\nabla\left(\frac{\zeta}{e}-\varphi\right)\right\}, \tag{4.7}$$

где g_p аналогично (4.3а).

Полный ток

$$j=j_n+j_p=ne\mu_n\left\{\nabla\left(\frac{\zeta}{e}-\varphi\right)+\frac{k_0}{e}\left(g_n-\frac{\zeta}{k_0T}\right)\nabla T\right\}+$$
$$+pe\mu_p\left\{\nabla\left(\frac{\zeta}{e}-\varphi\right)-\frac{k_0}{e}\left(g_p+\frac{\varepsilon_G+\zeta}{k_0T}\right)\nabla T\right\}. \tag{4.8}$$

Полагая полный ток $j=0$, получим для дифференциальной термоэдс собственного полупроводника

$$\alpha=\frac{\left|\nabla\left(\frac{\zeta}{e}-\varphi\right)\right|}{|\nabla T|}=\frac{k_0}{e}\frac{1}{n\mu_n+p\mu_p}\left\{n\mu_n\left(g_n-\frac{\zeta}{k_0T}\right)-\right.$$
$$\left.-p\mu_p\left(g_p+\frac{\varepsilon_G+\zeta}{k_0T}\right)\right\}=\frac{k_0}{e}\frac{1}{n\mu_n+p\mu_p}\left\{n\mu_n\left[g_n+\lg\frac{2\,(2\pi m_n k_0T)^{3/2}}{nh^3}\right]-\right.$$
$$\left.-p\mu_p\left(g_p+\lg\frac{2\,(2\pi m_p k_0T)^{3/2}}{ph^3}\right)\right\}, \tag{4.9}$$

где $-\zeta$ и $\varepsilon_G+\zeta$ взяты из (VI.2.15б), (6.2.15в). Мы видим, что вклады, вносимые в термоэдс электронами и дырками, имеют противоположные знаки; отсюда следует, что термоэдс собственного полупроводника, вообще говоря, меньше, чем примесного.

3. Для рассмотрения других термоэлектрических явлений вычислим плотность потока энергии, переносимого электронами

и дырками:

$$\boldsymbol{w}=\int f_1^{(n)}(\boldsymbol{k})\{\varepsilon-e\varphi\}\,\boldsymbol{v}_n\frac{d^3k}{4\pi^3}+$$
$$+\int f_1^{(p)}(\boldsymbol{k}')\{\varepsilon'+\varepsilon_G+e\varphi\}\,\boldsymbol{v}_p\frac{d^3k'}{4\pi^3}. \quad (4.10)$$

Индексы n и p, как всегда, отмечают величины, относящиеся к электронам и дыркам. Выражения, стоящие в скобках (4.10), учитывают наряду с кинетической и потенциальную энергию электронов и дырок; слагаемое ε_G учитывает, что энергия дырки ε' отсчитывается не от нижнего края зоны проводимости, а от верхнего края валентной зоны.

Неравновесные части функций распределения в потоке $\boldsymbol{w}$ с выражениями (4.2) и (4.6) для $\chi^*(\varepsilon)$ позволяют без труда вычислить $\boldsymbol{w}$. Те слагаемые в фигурных скобках в (4.10), которые не зависят от ε и ε', дают в $\boldsymbol{w}$ составляющие, пропорциональные $\boldsymbol{j}_n$ и $\boldsymbol{j}_p$. Слагаемые ε и ε' приводят к вычислению интегралов типа (2.20) от τx и τx^2. В результате получим[1])

$$\boldsymbol{w}=\varphi\boldsymbol{j}+\frac{\varepsilon_G}{e}\boldsymbol{j}_p-\frac{gk_0T}{e}n\mu_n\left[\nabla(\zeta-e\varphi)-\frac{\zeta-hk_0T}{T}\nabla T\right]+$$
$$+\frac{gk_0T}{e}p\mu_p\left[\nabla(\zeta-e\varphi)-\frac{\zeta+\varepsilon_G+hk_0T}{T}\nabla T\right], \quad (4.11)$$

где $\boldsymbol{j}_p$ выражается формулой (4.7) и

$$h=\langle\tau x^2\rangle/\langle\tau x\rangle. \quad (4.11а)$$

Подставим сюда (в том числе и в $\boldsymbol{j}_p$) $\nabla(\zeta-e\varphi)$ из (4.8). Приводя полученные слагаемые к общему знаменателю, объединяя подобные члены при $n\mu_n$, $p\mu_p$ и их произведениях, получим в результате длинных, но элементарных алгебраических вычислений

$$\boldsymbol{w}=\left(\varphi-\frac{\zeta}{e}\right)\boldsymbol{j}-\Pi\boldsymbol{j}-\varkappa\nabla T, \quad (4.12)$$

где коэффициент Пельтье

$$\Pi=T\frac{k_0}{e}\frac{1}{n\mu_n+p\mu_p}\left\{n\mu_n\left(g-\frac{\zeta}{k_0T}\right)-p\mu_p\left(g+\frac{\zeta+\varepsilon_G}{k_0T}\right)\right\} \quad (4.12а)$$

и коэффициент теплопроводности

$$\varkappa=n\mu_n(hg-g^2)\frac{k_0^2T}{e}+p\mu_p(hg-g^2)\frac{k_0^2T}{e}+$$
$$+\frac{n\mu_np\mu_p}{n\mu_n+p\mu_p}\frac{(\varepsilon_G+2gk_0T)^2}{eT}. \quad (4.12б)$$

Если время релаксации $\tau\sim\varepsilon^r$, то численное значение коэффициента в (4.12б)

$$hg-g^2=g^2(h-g)=r+\frac{5}{2}, \quad (4.12в)$$

[1]) Мы считаем, что $g_n=g_p\equiv g$ и $h_n=h_p\equiv h$, что, вообще говоря, обычно и имеет место.

т. е. совпадает с g (4.3б). Для разных механизмов рассеяния (3.5) этот коэффициент равен (4.4).

Из (4.12а) и (4.9) следует, что

$$\Pi = \alpha T. \tag{4.13}$$

Это так называемое *соотношение Томсона*, которое является, как мы сейчас покажем, следствием соотношений симметрии кинетических коэффициентов (1.5), поэтому оно справедливо не только для конкретной модели полупроводника, рассмотренного выше.

Полагая в (1.1) ток $\boldsymbol{j}=0$, получим для дифференциальной термоэдс (4.5)

$$\alpha = \frac{\left|\nabla\left(\frac{\zeta}{e}-\varphi\right)\right|}{|\nabla T|} = \frac{\beta}{\sigma}. \tag{4.14}$$

Подставляя $\nabla\left(\frac{\zeta}{e}-\varphi\right)$ из (1.1) в (1.2), получим

$$\boldsymbol{w} = \left(\varphi - \frac{\zeta}{e}\right)\boldsymbol{j} + \frac{\gamma}{\sigma}\boldsymbol{j} + \left(\gamma\frac{\beta}{\sigma} - \varkappa\right)\nabla T.$$

Сравнивая это выражение с (4.12), получим для коэффициента Пельтье

$$\Pi = \gamma/\sigma. \tag{4.15}$$

Из (1.5), записанного для скалярных коэффициентов γ и β (4.14) и (4.15), немедленно следует соотношение Томсона (4.13).

Из (4.12б) видно, что при смешанной проводимости, наряду с теплом, аддитивно переносимым электронами и дырками, часть теплопроводности обусловлена электронно-дырочными парами (слагаемое, пропорциональное произведению np в (4.12б)). Передача тепла осуществляется в этом случае за счет выделения энергии при рекомбинации электронов и дырок на холодном конце, где их равновесная концентрация ниже, чем в нагретой части полупроводника. Если концентрации и подвижности для электронов и дырок одного порядка и $\varepsilon_G \gg k_0 T$, то теплопроводность, обусловленная электронно-дырочными парами в $(\varepsilon_G/k_0T)^2$ раз больше теплопроводности электронов (дырок).

В полупроводниках нередко теплопроводность кристаллической решетки $\varkappa_0$ того же порядка, что и теплопроводность, обусловленная электронами (дырками). В этом случае мы должны правую часть (4.12) дополнить членом $-\varkappa_0\nabla T$.

Для того чтобы получить ряд интересных следствий из (4.12), возьмем дивергенцию от обеих частей (4.12) и учтем, что в стационарном состоянии $\operatorname{div}\boldsymbol{w} = \operatorname{div}\boldsymbol{j} = 0$; тогда

$$\operatorname{div}(-\varkappa\nabla T) = \boldsymbol{j}\left(\nabla\Pi + \nabla\left(\frac{\zeta}{e}-\varphi\right)\right). \tag{4.16}$$

При этом мы воспользовались тождеством $\operatorname{div}(\psi \boldsymbol{a}) = \psi \operatorname{div} \boldsymbol{a} + \boldsymbol{a} \nabla \psi$. Если применить выражение (4.16) к бесконечно малому цилиндру (рис. IX.2) с основаниями, параллельными границе между двумя проводниками, и воспользоваться теоремой Гаусса и интегральным преобразованием градиента [1]), то получим

$$-\varkappa_1 \left(\frac{\partial T}{\partial n}\right)_1 + \varkappa_2 \left(\frac{\partial T}{\partial n}\right)_2 = (\Pi_1 - \Pi_2)\, j. \qquad (4.17)$$

Здесь $\boldsymbol{n}$ — нормаль к границе, направленная из 1-го проводника во 2-й; предполагается, что $\boldsymbol{j}$ параллельно $\boldsymbol{n}$.

При выводе (4.17) необходимо учесть, что j и $\left(\frac{\zeta}{e} - \varphi\right)$ на границе непрерывны.

Если $\left(\frac{\partial T}{\partial n}\right)_1$ и $\left(\frac{\partial T}{\partial n}\right)_2$ имеют разные знаки, то слева в (4.17) стоит сумма двух тепловых потоков, текущих в противоположные стороны от границы (или к границе), а справа — количество тепла, выделяющегося (или поглощающегося) в 1 *сек* на 1 *см*^{2} границы.

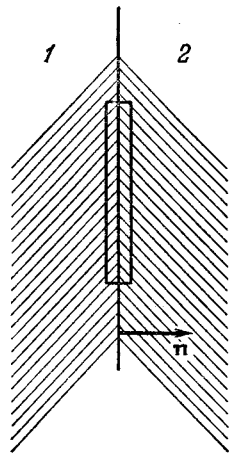


Рис. IX. 2.

В случае границы полупроводника с металлом можно для последнего пренебречь коэффициентом Пельтье.

Таким образом, выражение $(\Pi_1 - \Pi_2)\, j$ определяет выделение (поглощение) тепла за 1 *сек* на границе двух проводников при прохождении электрического тока.

Из (4.8) и (4.12а) найдем

$$\boldsymbol{j} = \sigma \nabla \left(\frac{\zeta}{e} - \varphi\right) + \frac{\sigma \Pi}{T} \nabla T. \qquad (4.18)$$

Подставляя отсюда $\nabla(\zeta/e - \varphi)$ в (4.16), получим

$$\operatorname{div}(-\varkappa \nabla T) = \frac{j^2}{\sigma} + \boldsymbol{j}\left(\nabla \Pi - \frac{\Pi}{T} \nabla T\right) = \frac{j^2}{\sigma} + \tau_T (\nabla T \boldsymbol{j}), \qquad (4.19)$$

где *коэффициент Томсона* τ_T может быть представлен в виде

$$\tau_T = T \frac{d}{dT}\left(\frac{\Pi}{T}\right) = T \frac{d\alpha}{dT}, \qquad (4.19а)$$

если воспользоваться соотношением (4.13).

Из (4.19) следует, что тепло, выделяющееся в 1 *сек* в единице объема (левая часть (4.19)), равно теплу Джоуля j^2/σ,

[1]) Кочин Н. Е., гл. II.

которое пропорционально квадрату силы тока, и теплу Томсона $\tau_T(\nabla T\boldsymbol{j})$ пропорциональному первой степени силы тока. Причем тепло Томсона, в зависимости от взаимной ориентации ∇T и $\boldsymbol{j}$, может как выделяться, так и поглощаться. Для примесного полупроводника (электронного или дырочного) коэффициент теплопроводности

$$\varkappa = n\mu\,(hg-g^2)\,\frac{k_0^2T}{e}\,. \tag{4.20}$$

Отношение коэффициента теплопроводности $\varkappa$ к удельной электропроводности $\sigma = en\mu$ равно

$$\frac{\varkappa}{\sigma} = (hg-g^2)\left(\frac{k_0}{e}\right)^2 T = LT, \tag{4.21}$$

т. е. пропорционально абсолютной температуре T (закон *Видемана—Франца*). При степенной зависимости τ от энергии ($\tau \backsim \varepsilon^r$) *число Лоренца*

$$L = (hg-g^2)\left(\frac{k_0}{e}\right)^2 = \left(r+\frac{5}{2}\right)\left(\frac{k_0}{e}\right)^2. \tag{4.21а}$$

Коэффициент $r+5/2$ в числе Лоренца имеет для разных механизмов рассеяния значения (4.4).

Пользуясь (4.12а), нетрудно написать выражение для коэффициента Пельтье Π в случае примесного (электронного или дырочного) полупроводника.

4. Термоэлектрические явления открывают техническую возможность непосредственного превращения тепловой энергии в электрическую (термоэдс) и возможность охлаждения посредством пропускания электрического тока через контакт двух проводников (эффект Пельтье).

Термоэдс металлов мала, в лучшем случае она составляет несколько десятков *мк/град*. Поэтому возможность технического использования термоэлементов в качестве эффективных термогенераторов электрического тока возникла только тогда, когда было установлено, что для полупроводников термоэдс достигает значений сотен *мк/град*.

Впервые А. Ф. Иоффе в 1929 г. указал на перспективность использования полупроводников для технических термогенераторов. Еще до войны А. Ф. Иоффе и его сотрудникам удалось получить термоэлементы с коэффициентом полезного действия (к.п.д.) порядка 3%.

Более широко развернулись под руководством А. Ф. Иоффе работы по техническому использованию термоэлектричества в послевоенный период. В настоящее время такие работы ведутся во всем мире, но главным образом в СССР, США и Англии.

В нашу задачу не входит изложение вопросов технического использования термоэлектрических явлений. Интересующихся этим вопросом мы отсылаем к соответствующей литературе [1]).

§ 5. Гальваномагнитные явления в невырожденных полупроводниках с простой зонной структурой

Гальваномагнитными эффектами называются явления, связанные с прохождением электрического тока при одновременном действии электрического поля $\boldsymbol{E}$ и магнитного поля $\boldsymbol{H}$.

Теория может быть построена в случае произвольного магнитного поля, однако в этом случае кинетические коэффициенты не выражаются через элементарные функции. Поэтому мы рассмотрим предельные случаи слабого и сильного магнитного поля. Мы будем называть магнитное поле слабым, если безразмерный параметр, входящий в (2.11), (2.12) (для простой зоны, когда $u = \hbar/m^*$)

$$\frac{e\tau u H}{\hbar c} = \frac{e\tau H}{cm^*} \approx \frac{\mu H}{c} \ll 1. \tag{5.1}$$

Здесь подвижность $\mu \approx e\tau/m^*$, где τ — соответствующим образом усредненное время релаксации (или просто время релаксации, если оно не зависит от энергии).

Критерий (5.1) может быть выражен в более наглядной форме. Учитывая, что $eH/cm^* = \omega_c$ — циклотронная частота (VI.6.1а), $v\tau = l$ — длина свободного пробега и $cm^*v/eH = R$ — радиус круговой орбиты свободного электрона в магнитном поле, получим

$$\omega_c \tau = \frac{l}{R} \ll 1. \tag{5.1а}$$

Таким образом, мы будем считать магнитное поле слабым, если длина свободного пробега электрона l много меньше радиуса его круговой орбиты в магнитном поле, или время свободного пробега τ много меньше $1/\omega_c = T_c/2\pi$, где T_c — период обращения электрона по этой орбите.

1. Рассмотрим вначале примесный полупроводник с одним сортом носителей тока, например электронами. Для упрощения записи у всех величин в выражениях (5.2)—(5.18) опущен

---

[1]) Иоффе А. Ф. Полупроводниковые термоэлементы — М.—Л.: Изд. АН СССР, 1960; Иоффе А. Ф., Стильбанс Л. С., Иорданишвили Е. К., Ставицкая Т. С. Термоэлектрическое охлаждение. — М.—Л.: Изд. АН СССР, 1959; Охотин А. С., Пушкарский А. С., Боровикова Р. П., Симонов В. А. Методы измерения характеристик термоэлектрических материалов и преобразователей. — М.: Наука, 1974.

индекс n. В изотермических условиях

$$\boldsymbol{P} = \nabla\left(\frac{\zeta}{e} - \varphi\right) = \boldsymbol{E}, \tag{5.2}$$

как и в случае (3.1).

Пусть магнитное поле $\boldsymbol{H}$ перпендикулярно электрическому полю $\boldsymbol{E}$, тогда, согласно (2.17),

$$\chi^* = \frac{\tau\boldsymbol{E} + \gamma\tau^2\,[\boldsymbol{H}\boldsymbol{E}]}{1 + (\gamma\tau H)^2}. \tag{5.3}$$

Направляя магнитное поле $\boldsymbol{H}$ по оси z, а электрическое в плоскости xy, получим из (5.3) и (2.21)

$$\begin{aligned} j_x &= a_1 E_x - a_2 E_y, \\ j_y &= a_2 E_x + a_1 E_y, \end{aligned} \tag{5.4}$$

где

$$a_1 = \frac{ne^2}{m}\left\langle\frac{\tau}{1 + (\gamma\tau H)^2}\right\rangle, \tag{5.4а}$$

$$a_2 = \frac{ne^2}{m}(\gamma H)\left\langle\frac{\tau^2}{1 + (\gamma\tau H)^2}\right\rangle. \tag{5.4б}$$

Эффект Холла заключается в возникновении электрического поля, перпендикулярного к току и магнитному полю, которое также перпендикулярно к току. Если ток течет вдоль оси x, так что $j_x = j$ и $j_y = 0$, то, исключая из уравнений (5.4) E_x, получим

$$E_y = -\frac{a_2}{a_1^2 + a_2^2}\, j = RHj, \tag{5.5}$$

где, по определению, R — постоянная Холла. Таким образом,

$$R = -\frac{a_2}{(a_1^2 + a_2^2)\,H}. \tag{5.5а}$$

В отсутствие магнитного поля ток $\boldsymbol{j} = \sigma\boldsymbol{E} = \boldsymbol{E}/\rho$, где σ и ρ — удельные электропроводность и сопротивление. При действии магнитного поля $\boldsymbol{H} \| z$ направление тока $\boldsymbol{j} \| x$ не совпадает с направлением электрического поля $\boldsymbol{E}$, поэтому удельное сопротивление в магнитном поле или *магнетосопротивление* $\rho_H = E_x/j = E_x/j_x$. Изменение сопротивления в магнитном поле

$$\frac{\rho_H - \rho}{\rho} \equiv \frac{\Delta\rho}{\rho} = \frac{E_x}{\rho j_x} - 1 = \frac{\sigma E_x}{j_x} - 1. \tag{5.6}$$

Полагая в уравнениях (5.4) $j_y = 0$ и определяя из них E_x, получим из (5.6)

$$\frac{\Delta\rho}{\rho} = \frac{\sigma a_1}{a_1^2 + a_2^2} - 1. \tag{5.6а}$$

Если τ от энергии не зависит, как это имеет место, например, в ионных полупроводниках при низких температурах, то

символ усреднения $\langle\rangle$ в (5.4а) и (5.4б) можно опустить. Легко показать, что в этом случае при любом магнитном поле

$$\left.\begin{aligned} R &= -1/cen, \\ \Delta\rho/\rho &= 0. \end{aligned}\right\} \tag{5.7}$$

Для τ, зависящего от энергии электрона ε, рассмотрим случаи слабого и сильного магнитных полей.

А. Слабые магнитные поля: $\gamma\tau H \ll 1$ (условие (5.1)). В линейном приближении по магнитному полю, т. е., удерживая в разложении в ряд по H члены не выше первой степени, получим

$$a_1 = \frac{ne^2}{m}\langle\tau\rangle, \quad a_2 = \frac{ne^2}{m}(\gamma H)\langle\tau^2\rangle. \tag{5.8}$$

В квадратичном приближении, т. е., удерживая члены порядка H^2, имеем

$$a_1 = \frac{ne^2}{m}\{\langle\tau\rangle - (\gamma H)^2\langle\tau^3\rangle\}, \quad a_2 = \frac{ne^2}{m}(\gamma H)\langle\tau^2\rangle. \tag{5.9}$$

Из (5.5а) и (5.6а) следует, что в линейном приближении по H для R и в квадратичном[1]) для $\Delta\rho/\rho$ $\left(\text{в линейном приближении по } H\ \frac{\Delta\rho}{\rho} = 0\right)$

$$R = -\frac{\langle\tau^2\rangle}{\langle\tau\rangle^2}\frac{1}{ecn}, \tag{5.10}$$

$$\frac{\Delta\rho}{\rho} = \left(\frac{eH}{mc}\right)^2 \frac{\langle\tau^3\rangle\langle\tau\rangle - \langle\tau^2\rangle^2}{\langle\tau\rangle^2}. \tag{5.11}$$

Из этих выражений непосредственно видно, что если τ от энергии не зависит, то R и $\Delta\rho/\rho$ равны выражениям, указанным в (5.7).

Если τ зависит от энергии по степенному закону (3.4), то, используя определение (2.20), легко найти численные коэффициенты в (5.10) и (5.11). Для разных механизмов рассеяния (3.5), получим

$$\text{а)}\ R = -\frac{3\pi}{8}\frac{1}{ecn}, \quad \frac{\Delta\rho}{\rho} = \left(\frac{eH}{mc}\right)^2 \frac{\tau_{0L}^2}{(k_0T)^3}\left[1 - \frac{\pi}{4}\right] = \frac{9\pi}{16}\left[1 - \frac{\pi}{4}\right]\left(\frac{\mu H}{c}\right)^2, \tag{5.12а}$$

---

[1]) При вычислении $\Delta\rho/\rho$ надо воспользоваться разложением $\frac{1}{a + bH^2} = \frac{1}{a}\frac{1}{1 + \frac{b}{a}H^2} \approx \frac{1}{a}\left(1 - \frac{b}{a}H^2\right)$.

б) $$R=-\frac{315\pi}{512}\frac{1}{ecn},\qquad \frac{\Delta\rho}{\rho}=\left(\frac{eH}{mc}\right)^2\tau_{0l}^2(k_0T)^3\,120\left[1-\frac{6615\pi}{32\,768}\right]=$$
$$=\frac{15\pi}{8}\left[1-\frac{6615\pi}{32\,768}\right]\left(\frac{\mu H}{c}\right)^2, \quad (5.12б)$$

в) $$R=-\frac{45\pi}{128}\frac{1}{ecn},\ \frac{\Delta\rho}{\rho}=\left(\frac{eH}{mc}\right)^2\frac{\tau_{0\rho}^2}{(k_0T)}3\left[1-\frac{75\pi}{256}\right]=$$
$$=\frac{27\pi}{64}\left[1-\frac{75\pi}{256}\right]=\left(\frac{\mu H}{c}\right)^2, \quad (5.12в)$$

г) см. (5.7).

Отношение $\Delta\rho/\rho$ мы выразили через подвижности (3.6а)—(3.6в). Мы видим, что в слабых полях $\Delta\rho/\rho\sim H^2$. Из (5.10) следует, что

$$|R|\sigma=|R|ne\mu=\frac{\langle\tau^2\rangle}{\langle\tau\rangle^2}\frac{1}{c}\mu, \quad (5.13)$$

где σ—удельная электропроводность, μ—подвижность. Таким образом, в случае носителей одного знака измерение R позволяет определить их концентрацию, а измерение R и σ—их подвижность.

Введем понятие об *угле Холла* θ между направлением тока $\boldsymbol{j}$ и поля $\boldsymbol{E}$:

$$\operatorname{tg}\theta\approx\theta=\left|\frac{E_y}{E_x}\right|=\left|\frac{a_2}{a_1}\right|=\frac{\langle\tau^2\rangle}{\langle\tau\rangle}\frac{eH}{mc}=\frac{\langle\tau^2\rangle}{\langle\tau\rangle^2}\frac{\mu H}{c}, \quad (5.14)$$

откуда видно, что $\theta\approx\frac{\mu H}{c}\ll 1$.

Б. Сильные магнитные поля: $\gamma\tau H\gg 1$. Пренебрегая единицей по сравнению с $(\gamma\tau H)^2$, получим из (5.4а), (5.4б)

$$a_1=\frac{ne^2}{m}\frac{1}{(\gamma H^2)}\left\langle\frac{1}{\tau}\right\rangle;\quad a_2=\frac{ne^2}{m}\frac{1}{\gamma H}. \quad (5.15)$$

При этом $\frac{a_1}{a_2}=\left\langle\frac{1}{\tau}\right\rangle\frac{1}{\gamma H}\ll 1$, так что $a_1^2+a_2^2\approx a_2^2$. Из (5.5а) и (5.6а) получим, используя (5.15),

$$R=-1/ecn, \quad (5.16)$$
$$\frac{\Delta\rho}{\rho}=\langle\tau\rangle\left\langle\frac{1}{\tau}\right\rangle-1. \quad (5.17)$$

Мы видим, что в случае сильных магнитных полей постоянная Холла не зависит от механизма рассеяния, а магнетосопротивление $\Delta\rho/\rho$ достигает насыщения.

Для различных механизмов рассеяния (3.5) имеем

$$\text{а)}\ \frac{\Delta\rho}{\rho}=\frac{32}{9\pi}-1,\quad \text{б)}\ \frac{\Delta\rho}{\rho}=\frac{32}{3\pi}-1,$$
$$\text{в)}\ \frac{\Delta\rho}{\rho}=\frac{32}{9\pi}-1,\quad \text{г)}\ \frac{\Delta\rho}{\rho}=0. \quad (5.18)$$

Заметим, что если магнитное поле $\boldsymbol{H}$ параллельно электрическому $\boldsymbol{E}$, то в полупроводниках со стандартной зоной, магнитное поле не оказывает влияния на явления переноса.

В самом деле, в случае $\boldsymbol{H} \parallel \boldsymbol{E}$ ($\boldsymbol{E} = \boldsymbol{P}$) из (2.17) следует

$$\boldsymbol{\chi}^* = \frac{\tau \boldsymbol{E} + \gamma^2 \tau^3 H^2 \boldsymbol{E}}{1 + (\gamma \tau H)^2} = \tau \boldsymbol{E},$$

откуда видно, что магнитное поле не влияет на ток.

Таким образом, в этом случае эффект Холла и магнетосопротивление равны нулю.

2. Рассмотрим теперь гальваномагнитные явления в полупроводниках со смешанной проводимостью, когда имеются и электроны и дырки.

В изотермических условиях ($\nabla T = 0$) $\boldsymbol{P}_n$ и $\boldsymbol{P}_p$ равны $\boldsymbol{E}$. Если $\boldsymbol{H}$ по-прежнему перпендикулярно $\boldsymbol{E}$, то $\boldsymbol{\chi}_n^*$ равно (5.3)

$$\boldsymbol{\chi}_n^* = \frac{\tau_n \boldsymbol{E} + \gamma_n \tau_n^2 [\boldsymbol{HE}]}{1 + (\gamma_n \tau_n H)^2}, \tag{5.19}$$

а $\boldsymbol{\chi}_p^*$, как следует из (2.23), равно

$$\boldsymbol{\chi}_p^* = \frac{\tau_p \boldsymbol{E} - \gamma_p \tau_p^2 [\boldsymbol{HE}]}{1 + (\gamma_p \tau_p H)^2}. \tag{5.19а}$$

Здесь γ_n и γ_p равны (2.17б) и (2.23б). Плотность электрического тока, как следует из (2.25), равна

$$\boldsymbol{j} = \boldsymbol{j}_n + \boldsymbol{j}_p = \frac{ne^2}{m_n} \langle \boldsymbol{\chi}_n^* \rangle + \frac{pe^2}{m_p} \langle \boldsymbol{\chi}_p^* \rangle. \tag{5.20}$$

Направляя магнитное поле $\boldsymbol{H}$ по оси z, а электрическое в плоскости xy, получим из (5.20), (5.19) и (5.19а) уравнения (5.4), в которых

$$a_1 = \frac{ne^2}{m_n} \left\langle \frac{\tau_n}{1 + (\gamma_n \tau_n H)^2} \right\rangle + \frac{pe^2}{m_p} \left\langle \frac{\tau_p}{1 + (\gamma_p \tau_p H)^2} \right\rangle, \tag{5.21}$$

$$a_2 = \frac{ne^2}{m_n} (\gamma_n H) \left\langle \frac{\tau_n^2}{1 + (\gamma_n \tau_n H)^2} \right\rangle - \frac{pe^2}{m_p} (\gamma_p H) \left\langle \frac{\tau_p^2}{1 + (\gamma_p \tau_p H)^2} \right\rangle. \tag{5.21а}$$

Мы можем, аналогично предыдущему пункту, рассмотреть случаи слабого и сильного магнитных полей.

А. Слабое магнитное поле ($\gamma_n \tau_n H \ll 1$ и $\gamma_p \tau_p H \ll 1$). Определяя a_1, a_2 и отношение $a_2/(a_1^2 + a_2^2)$ с точностью до первой степени магнитного поля H, получим из (5.5а) для постоянной Холла

$$R = -\frac{a_2}{(a_1^2 + a_2^2) H} = -\frac{1}{ce} \frac{n \frac{e^2}{m_n^2} \langle \tau_n^2 \rangle - p \frac{e^2}{m_p^2} \langle \tau_p^2 \rangle}{\left[n \frac{e}{m_n} \langle \tau_n \rangle + p \frac{e}{m_p} \langle \tau_p \rangle \right]^2}. \tag{5.22}$$

В случае примесного полупроводника $p = 0$ (или $n = 0$) и (5.22) превращается в (5.10).

Используя (3.3), можно выразить R через подвижности:

$$R=-\frac{1}{ce}\frac{\dfrac{\langle\tau_n^2\rangle}{\langle\tau_n\rangle^2}n\mu_n^2-\dfrac{\langle\tau_p^2\rangle}{\langle\tau_p\rangle^2}p\mu_p^2}{[n\mu_n+p\mu_p]^2}. \tag{5.23}$$

Если, например, электроны и дырки рассеиваются на акустических колебаниях (3.5а), то

$$\frac{\langle\tau_n^2\rangle}{\langle\tau_n\rangle^2}=\frac{\langle\tau_p^2\rangle}{\langle\tau_p\rangle^2}=\frac{8\pi}{3} \tag{5.23а}$$

и

$$R=-\frac{8\pi}{3ce}\frac{n\mu_n^2-p\mu_p^2}{(n\mu_n+p\mu_p)^2}. \tag{5.23б}$$

Если для собственного полупроводника $n=p$, то

$$R=-\frac{8\pi}{3cen}\frac{\mu_n^2-\mu_p^2}{(\mu_n+\mu_p)^2}. \tag{5.23в}$$

Отсюда видно, что если μ_n мало отличается от μ_p, постоянная Холла R мала.

Так как магнетосопротивление $\Delta\rho/\rho$ в линейном приближении по магнитному полю равно нулю, то необходимо определить a_1 и a_2 и вычислить $\Delta\rho/\rho$ (5.6а) в квадратичном приближении по магнитному полю; в результате довольно длинных, но элементарных вычислений получим

$$\frac{\Delta\rho}{\rho}=\frac{1}{\sigma^2}\left\{\left(\frac{ne^2}{m_n}\right)^2(\gamma_nH)^2[\langle\tau_n\rangle\langle\tau_n^3\rangle-\langle\tau_n^2\rangle^2]+\right.$$
$$+\left(\frac{pe^2}{m_p}\right)^2[(\gamma_pH)^2\langle\tau_p\rangle\langle\tau_p^3\rangle-\langle\tau_p^2\rangle^2]+$$
$$\left.+\left(\frac{ne^2}{m_n}\right)\left(\frac{pe^2}{m_p}\right)[(\gamma_pH)^2\langle\tau_n\rangle\langle\tau_p^3\rangle+(\gamma_nH)^2\langle\tau_p\rangle\langle\tau_n^3\rangle]\right\}. \tag{5.24}$$

Отметим, что в слабом магнитном поле $\Delta\rho/\rho\sim H^2$. Для примесного полупроводника ($p=0$) из (5.24) получим (5.11).

Используя (3.3), можно выразить $\Delta\rho/\rho$ через подьижности μ_n и μ_p аналогично (5.23). Не выписывая общей формулы, укажем, что в случае рассеяния на акустических колебаниях

$$\frac{\Delta\rho}{\rho}=\frac{9\pi e}{16c^2}\frac{H^2}{\sigma}\left[(n\mu_n^3+p\mu_p^3)-\frac{\pi}{4}\frac{(n\mu_n^2-p\mu_p^2)^2}{n\mu_n+p\mu_p}\right]. \tag{5.25}$$

Для примесного полупроводника ($p=0$, $\sigma=en\mu_n$) это выражение переходит в (5.12а).

Б. Сильное магнитное поле ($\gamma_n\tau_nH\gg1$ и $\gamma_p\tau_pH\gg1$). В этом случае разложение ведется по обратным степеням магнитного поля $1/H$. В приближении $1/H^2$ из (5.21) и (5.21а) следует

$$a_1=nm_n\frac{c^2}{H^2}\left\langle\frac{1}{\tau_n}\right\rangle+pm_p\frac{c^2}{H^2}\left\langle\frac{1}{\tau_p}\right\rangle, \tag{5.26}$$

$$a_2=(n-p)\frac{ec}{H}. \tag{5.26а}$$

Если $n \neq p$, то $\frac{a_1}{a_2} \approx \frac{1}{\gamma H}\left\langle\frac{1}{\tau}\right\rangle \ll 1$, поэтому $a_1^2 + a_2^2 \approx a_2^2$. Постоянная Холла (5.5а)

$$R = -\frac{a_2}{(a_1^2+a_2^2)\,H} = -\frac{1}{a_2 H} = -\frac{1}{ec\,(n-p)}. \tag{5.27}$$

Магнетосопротивление (5.6а) равно

$$\frac{\Delta\rho}{\rho} = \frac{\sigma a_1}{a_2} - 1 = \frac{\sigma\left[nm_n\left\langle\frac{1}{\tau_n}\right\rangle + pm_p\left\langle\frac{1}{\tau_p}\right\rangle\right]}{e^2\,(n-p)^2} - 1, \tag{5.28}$$

т. е. достигает насыщения; при $p=0$ и $\sigma = (e^2 n/m_n)\langle\tau_n\rangle$ (5.28) превращается в (5.17).

Случай $n=p$ требует особого рассмотрения. На первый взгляд может показаться, что из-за того, что в этом случае $a_2 = 0$ (5.26а), а $a_1 \neq 0$ (5.26), постоянная Холла $R=0$. Однако в приближении $1/H^2$ не только $a_2 = 0$, но и $(a_1^2+a_2^2)\,H \sim 1/H^3 \approx 0$; поэтому для определения R надо вычислить a_2 и $(a_1^2+a_2^2)\,H$ с точностью до $1/H^3$. В этом случае

$$a_2 = -\frac{nc^3}{eH^3}\left[m_n^2\left\langle\frac{1}{\tau_n^2}\right\rangle - m_p^2\left\langle\frac{1}{\tau_p^2}\right\rangle\right]. \tag{5.29}$$

Используя это значение a_2 и a_1 (5.26), получим для константы Холла в приближении $1/H^3$

$$R = -\frac{a_2}{(a_1^2+a_2^2)\,H} = -\frac{1}{ecn}\,\frac{m_p^2\left\langle\frac{1}{\tau_p^2}\right\rangle - m_n^2\left\langle\frac{1}{\tau_n^2}\right\rangle}{\left[m_p\left\langle\frac{1}{\tau_p}\right\rangle + m_n\left\langle\frac{1}{\tau_n}\right\rangle\right]^2}. \tag{5.30}$$

Выразим m_n и m_p из соотношения $\mu = \frac{e}{m}\langle\tau\rangle$ через подвижности μ_n и μ_p. Если электроны и дырки рассеиваются по одинаковому закону, то $\langle\tau_n\rangle\left\langle\frac{1}{\tau_n}\right\rangle = \langle\tau_p\rangle\left\langle\frac{1}{\tau_p}\right\rangle$ и $\langle\tau_n\rangle^2\left\langle\frac{1}{\tau_n^2}\right\rangle = \langle\tau_p\rangle^2\left\langle\frac{1}{\tau_p^2}\right\rangle$; тогда

$$R = -\frac{\langle 1/\tau_n^2\rangle}{\langle 1/\tau_n\rangle^2}\,\frac{1}{ecn}\,\frac{1-b}{1+b}, \tag{5.30а}$$

где $b = \mu_p/\mu_n$. Если τ_n и τ_p от энергии не зависят, то численный коэффициент в (5.30а) равен единице. При рассеянии на акустических колебаниях $\langle 1/\tau_n^2\rangle\langle 1/\tau_n\rangle^2 = 45\pi/128$. Магнетосопротивление (5.6а) в сильном магнитном поле при $n=p$ можно вычислить в приближении $1/H^2$, когда $a_2 = 0$, а a_1 равно (5.26)

$$\frac{\Delta\rho}{\rho} = \frac{\sigma}{a_1} - 1 = \frac{\sigma H^2}{c^2 n\,[m_n\langle 1/\tau_n\rangle + m_p\langle 1/\tau_p\rangle]} - 1 =$$

$$= \left(\frac{eH}{c}\right)^2 \frac{\frac{1}{m_n}\langle\tau_n\rangle + \frac{1}{m_p}\langle\tau_p\rangle}{m_n\langle 1/\tau_n\rangle + m_p\langle 1/\tau_p\rangle}. \tag{5.31}$$

Здесь мы пренебрегли единицей по сравнению с первым слагаемым и положили $\sigma = ne^2\left[\frac{1}{m_n}\langle\tau_n\rangle + \frac{1}{m_p}\langle\tau_p\rangle\right]$.

Заметим, что при $n = p$ в сильном магнитном поле магнетосопротивление $\Delta\rho/\rho \sim H^2$, так же как в слабом поле (5.24). Если электроны и дырки рассеиваются по одинаковому закону, мы можем поступить так же, как при переходе от (5.30) к (5.30а), и получим

$$\frac{\Delta\rho}{\rho} = \frac{1}{\langle\tau_n\rangle\langle 1/\tau_n\rangle}\left(\frac{\sqrt{\mu_n\mu_p}H}{c}\right)^2. \tag{5.31а}$$

В случае рассеяния электронов и дырок на акустических колебаниях численный коэффициент в (5.31а) равен $9\pi/32$.

§ 6. Термомагнитные явления в невырожденных полупроводниках с простой зонной структурой

1. Если наряду с электрическим и магнитным полями в полупроводнике существует градиент температуры, то явления переноса в нем называются *термомагнитными*.

Рассмотрим вначале примесный полупроводник с одним сортом носителей тока, например электронами. В этом случае (опуская для простоты записи индекс n), как следует из (2.17а),

$$\boldsymbol{P} = \frac{\varepsilon-\zeta}{eT}\nabla T + \nabla\left(\frac{\zeta}{e} - \varphi\right). \tag{6.1}$$

Если магнитное поле $\boldsymbol{H}$ перпендикулярно вектору $\boldsymbol{P}$, то, как следует из (2.17),

$$\boldsymbol{\chi}^* = \frac{\tau\boldsymbol{P} + \gamma\tau^2[\boldsymbol{HP}]}{1+(\gamma\tau H)^2}. \tag{6.2}$$

Направим магнитное поле $\boldsymbol{H}$ по оси z, а векторы ∇T и $\nabla\left(\frac{\zeta}{e}-\varphi\right)$ — в плоскости xy. Составляющие тока по осям x и y, аналогично (5.4), равны

$$\begin{aligned} j_x &= a_1\nabla_x\left(\frac{\zeta}{e}-\varphi\right) + b_1\nabla_x T - a_2\nabla_y\left(\frac{\zeta}{e}-\varphi\right) - b_2\nabla_y T,\\ j_y &= a_2\nabla_x\left(\frac{\zeta}{e}-\varphi\right) + b_2\nabla_x T + a_1\nabla_y\left(\frac{\zeta}{e}-\varphi\right) + b_1\nabla_y T, \end{aligned} \tag{6.3}$$

где a_1 и a_2 равны (5.4а), (5.4б), а

$$b_1 = \frac{ne^2}{m}\left\langle\frac{\left[\left(\frac{k_0}{e}\right)x - \frac{\zeta}{eT}\right]\tau}{1+(\gamma\tau H)^2}\right\rangle, \tag{6.3а}$$

$$b_2 = \frac{ne^2}{m}(\gamma H)\left\langle\frac{\left[\left(\frac{k_0}{e}\right)x - \frac{\zeta}{eT}\right]\tau^2}{1+(\gamma\tau H)^2}\right\rangle. \tag{6.3б}$$

Здесь, как обычно, $x = \varepsilon/k_0T$.

В качестве первого примера термомагнитных явлений рассмотрим эффект *Нернста—Эттингсгаузена*[1]), который заключается в возникновении поперечного поля $\nabla_y\left(\frac{\zeta}{e}-\varphi\right)$, при токе $\boldsymbol{j}=0$, но $\nabla_x T\neq 0$; рассмотрим вначале изотермический эффект Нернста, для которого $\nabla_y T=0$.

Полагая в уравнениях (6.3) $j_x=j_y=\nabla_y T=0$ и исключая из них $\nabla_x\left(\frac{\zeta}{e}-\varphi\right)$, получим

$$\nabla_y\left(\frac{\zeta}{e}-\varphi\right)=E_y=-\frac{a_1b_2-a_2b_1}{(a_1^2+a_2^2)}\nabla_x T=-QH\nabla_x T, \tag{6.4}$$

где *постоянная Нернста*

$$Q=\frac{a_1b_2-a_2b_1}{(a_1^2+a_2^2)H}. \tag{6.4а}$$

Из выражений для коэффициентов a_i и b_i (5.4а), (5.4б) и (6.3а), (6.3б) следует, что если время релаксации τ от энергии электрона ε не зависит, то $Q=0$. Как мы увидим ниже, знак Q зависит от характера зависимости τ от энергии ε.

Рассмотрим теперь эффект Нернста в слабых и сильных магнитных полях.

А. Слабые поля: $\gamma\tau H\ll 1$. В линейном приближении по магнитному полю коэффициенты a_1 и a_2 равны формулам (5.8), а коэффициенты b_1 и b_2, согласно (6.3а), (6.3б), равны

$$\begin{aligned} b_1&=\frac{ne^2}{m}\left\{\left(\frac{k_0}{e}\right)\langle\tau x\rangle-\frac{\zeta}{eT}\langle\tau\rangle\right\},\\ b_2&=\frac{ne^2}{m}(\gamma H)\left\{\left(\frac{k_0}{e}\right)\langle\tau^2 x\rangle-\frac{\zeta}{eT}\langle\tau^2\rangle\right\}, \end{aligned} \tag{6.5}$$

откуда

$$Q=\left(\frac{k_0}{e}\right)\frac{e}{mc}\frac{\langle\tau\rangle\langle\tau^2 x\rangle-\langle\tau^2\rangle\langle\tau x\rangle}{\langle\tau\rangle^2}. \tag{6.6}$$

Если τ имеет вид (3.4), то (см. Приложение 7)

$$Q=\left(\frac{k_0}{e}\right)\frac{e}{mc}\tau_0\frac{\Gamma\left(2r+\frac{5}{2}\right)}{\Gamma\left(r+\frac{5}{2}\right)}=\frac{3\sqrt{\pi}}{4c}=\frac{\Gamma\left(2r+\frac{5}{2}\right)}{\Gamma^2\left(r+\frac{5}{2}\right)}r\left(\frac{k_0}{e}\right)\mu. \tag{6.6а}$$

Мы можем определить теперь постоянную Нернста для различных механизмов рассеяния (3.5):

а) $$Q=-\frac{\sqrt{\pi}}{4}\left(\frac{k_0}{e}\right)\frac{e}{mc}\frac{a_A}{(k_0T)^{3/2}}=-\frac{3\pi}{16c}\left(\frac{k_0}{e}\right)\mu, \tag{6.7а}$$

б) $$Q=\frac{945\pi}{128}\left(\frac{k_0}{e}\right)\frac{e}{mc}a_I(k_0T)^{3/2}=\frac{945\pi}{1024c}\left(\frac{k_0}{e}\right)\mu, \tag{6.7б}$$

[1]) Для краткости мы будем называть его в дальнейшем *эффектом Нернста*.

в) $$Q=\frac{15\sqrt{\pi}}{32}\left(\frac{k_0}{e}\right)\frac{e}{mc}\frac{a_P}{(k_0T)^{1/2}}=\frac{45\pi}{256c}\left(\frac{k_0}{e}\right)\mu, \tag{6.7в}$$

г) $$Q=0. \tag{6.7г}$$

Знак Q совпадает со знаком показателя степени r в (3.4). Поэтому, если с изменением температуры один механизм рассеяния сменяется другим, то Q может при этом изменить знак.

Б. Сильные поля: $\gamma\tau H\gg 1$. Пренебрегая единицей по сравнению с $(\gamma\tau H)^2$, получим из (6.3а), (6.3б)

$$\begin{aligned} b_1&=\frac{ne^2}{m}\frac{1}{(\gamma H)^2}\left\{\left(\frac{k_0}{e}\right)\left\langle\frac{x}{\tau}\right\rangle-\frac{\zeta}{eT}\left\langle\frac{1}{\tau}\right\rangle\right\},\\ b_2&=\frac{ne^2}{m}\frac{1}{(\gamma H)}\left\{\left(\frac{k_0}{e}\right)\langle x\rangle-\frac{\zeta}{eT}\right\}. \end{aligned} \tag{6.8}$$

Подставляя (6.8) и (5.15) в (6.4а), получим в приближении $1/H^2$

$$Q=\frac{1}{\gamma H^2}\left(\frac{k_0}{e}\right)\left\{\langle x\rangle\left\langle\frac{1}{\tau}\right\rangle-\left\langle\frac{x}{\tau}\right\rangle\right\}. \tag{6.9}$$

Заметим, что из (3.4а) следует, что $\langle x\rangle=5/2$.

Легко показать, что в случае рассеяния электронов на акустических колебаниях (3.5а)

$$Q=-\frac{4}{3\sqrt{\pi}}\left(\frac{k_0}{e}\right)\frac{mc}{eH^2}\frac{(k_0T)^{3/2}}{a_A}=-\frac{16}{9\pi}\left(\frac{k_0}{e}\right)\frac{c}{\mu H^2}. \tag{6.10}$$

Аналогично определяется Q для других механизмов рассеяния.

Рассмотрим теперь изменение термоэдс в магнитном поле. Полагая опять в уравнениях (6.3) $j_x=j_y=0$, $\nabla_y T=0$, но исключая теперь $\nabla_y\left(\frac{\zeta}{e}-\varphi\right)$, получим для дифференциальной термоэдс в магнитном поле

$$\alpha_H=\frac{|\nabla_x(\varphi-\zeta/e)|}{|\nabla_x T|}=\frac{a_1b_1+a_2b_2}{a_1^2+a_2^2}. \tag{6.11}$$

При $H=0$, $a_2=b_2=0$ и термоэдс

$$\alpha_{H=0}=\frac{b_1}{a_1}\Big|_{H=0}=\frac{k_0}{e}\left\{\frac{\langle\tau x\rangle}{\langle\tau\rangle}-\frac{\zeta}{k_0T}\right\}=\alpha \tag{6.11а}$$

совпадает с (4.5).

Легко показать, что в линейном приближении по магнитному полю изменение термоэдс $\Delta\alpha=\alpha_H-\alpha=0$. Таким образом, для вычисления $\Delta\alpha$ надо определить b_1 и b_2 в квадратичном приближении по магнитному полю. Из (6.3а), (6.3б) следует

$$b_1=\frac{ne^2}{m}\left\{\left(\frac{k_0}{e}\right)\langle\tau x\rangle-\frac{\zeta}{eT}\langle\tau\rangle-(\gamma H)^2\left[\left(\frac{k_0}{e}\right)\langle\tau^3x\rangle-\frac{\zeta}{eT}\langle\tau^3\rangle\right]\right\}, \tag{6.12}$$

$$b_2=\frac{ne^2}{m}(\gamma H)\left\{\left(\frac{k_0}{e}\right)\langle\tau^2x\rangle-\frac{\zeta}{eT}\langle\tau^2\rangle\right\}. \tag{6.12а}$$

Из (6.12), (6.12а), (5.9), (6.11) и используя примечание на стр. 515, получим в квадратичном приближении по магнитному полю

$$\Delta\alpha = \left(\frac{k_0}{e}\right)(\gamma H)^2 \frac{\langle\tau\rangle\langle\tau^3\rangle\langle\tau x\rangle + \langle\tau\rangle\langle\tau^2\rangle\langle\tau^2 x\rangle - \langle\tau\rangle^2\langle\tau^3 x\rangle - \langle\tau^2\rangle^2\langle\tau x\rangle}{\langle\tau\rangle^3}. \tag{6.13}$$

Если τ не зависит от энергии электрона, то $\Delta\alpha = 0$ (это можно показать и для случая произвольного магнитного поля).

Для взаимодействия с акустическими колебаниями

$$\Delta\alpha = \left(\frac{k_0}{e}\right)\left(\frac{eH}{mc}\right)^2 \frac{\tau_A^2}{(k_0T)^3}\left(1 - \frac{\pi}{8}\right) = \frac{9\pi}{16}\left(1 - \frac{\pi}{8}\right)\left(\frac{k_0}{e}\right)\left(\frac{\mu H}{c}\right)^2. \tag{6.14}$$

При рассеянии на ионах примеси

$$\Delta\alpha = -\left(\frac{k_0}{e}\right)\left(\frac{eH}{mc}\right)^2 \tau_I^2 (k_0T)^3\, 340\left[1 - \frac{297\,675\pi}{2\,785\,280}\right] \approx -50{,}0\left(\frac{k_0}{e}\right)\left(\frac{\mu H}{c}\right)^2. \tag{6.15}$$

Следует отметить, что в случае (6.14) $\Delta\alpha$ возрастает, а в случае (6.15) — уменьшается при увеличении магнитного поля.

В случае сильных магнитных полей, когда можно пренебречь единицей по сравнению с $(\gamma\tau H)^2$, коэффициенты a_i и b_i равны выражениям (5.15) и (6.8). Подставляя эти значения в (6.11) и пренебрегая величинами $1/H^4$ по сравнению с $1/H^2$ и $1/H^2$ по сравнению с членами, не зависящими от H обнаружим, что α при $H \to \infty$ достигает насыщения:

$$\alpha_\infty = \left(\frac{k_0}{e}\right)\left\{\frac{5}{2} - \frac{\zeta}{k_0T}\right\}. \tag{6.16}$$

Наиболее интересной особенностью α_∞ является то, что она не зависит от механизма рассеяния. Отсюда и из (4.5) следует

$$\Delta\alpha = \alpha_\infty - \alpha = \left(\frac{k_0}{e}\right)\left\{\frac{5}{2} - \frac{\langle\tau x\rangle}{\langle\tau\rangle}\right\}. \tag{6.17}$$

Для взаимодействия с акустическими колебаниями и для рассеяния на ионах примеси наибольшие изменения α равны

$$\Delta\alpha = \frac{1}{2}\left(\frac{k_0}{e}\right) \text{ и } \Delta\alpha = -\frac{3}{2}\left(\frac{k_0}{e}\right). \tag{6.17а}$$

Определим еще изменение электронной теплопроводности в поперечном магнитном поле.

Поток тепловой энергии, аналогично (4.10), равен

$$\boldsymbol{w} = -\frac{en}{m}\langle\boldsymbol{\chi}^*(\varepsilon - e\varphi)\rangle = \varphi\frac{ne^2}{m}\langle\boldsymbol{\chi}^*\rangle - \frac{en}{m}\langle\boldsymbol{\chi}^*\varepsilon\rangle = \varphi\boldsymbol{j} - \frac{en}{m}k_0T\langle\boldsymbol{\chi}^* x\rangle, \tag{6.18}$$

где $\boldsymbol{\chi}^*$ равно (6.2); мы воспользовались тем, что поток электронов равен $\boldsymbol{j}/(-e)$, где электрический ток $\boldsymbol{j}$ равен (2.21).

Если градиент температуры направлен вдоль оси x, то $\nabla_y T = 0$; однако из-за наличия магнитного поля не только w_x, но и w_y отлично от нуля.

Полагая в уравнениях (6.3) $j_x = j_y = \nabla_y T = 0$, определим из них $\nabla_x\left(\frac{\zeta}{e} - \varphi\right)$ и $\Delta_y\left(\frac{\zeta}{e} - \varphi\right)$ через $\nabla_x T$ и подставим их в составляющую теплового потока w_x, которая будет тогда пропорциональна $\nabla_x T$. Отношение $w_x/(-\nabla_x T)$ равно коэффициенту теплопроводности электронов в магнитном поле $\varkappa(H)$. В случае слабого и сильного магнитного поля теплопроводность может быть выражена через элементарные функции. Не приводя эти элементарные, но громоздкие вычисления, укажем только результат для слабого поля в случае когда время релаксации определяется выражением (3.4):

$$\varkappa(H) = n\mu\left(r + \frac{5}{2}\right)\frac{k_0^2 T}{e}\left\{1 - a_r\left(\frac{\mu H}{c}\right)^2\right\}, \qquad (6.19)$$

где

$$a_r = \frac{9\pi}{16\left(r+\frac{5}{2}\right)}\left[\left(4r^2+3r+\frac{5}{2}\right)\frac{\Gamma\left(3r+\frac{5}{2}\right)}{\Gamma^3\left(r+\frac{5}{2}\right)} - r^2\frac{\Gamma^2\left(2r+\frac{5}{2}\right)}{\Gamma^4\left(r+\frac{5}{2}\right)}\right]. \qquad (6.19а)$$

Мы видим, что поправка к теплопроводности в слабом магнитном поле пропорциональна H^2.

При рассеянии на акустических колебаниях $r = -1/2$ поэтому

$$\varkappa(H) = n\mu\frac{2k_0^2 T}{e}\left\{1 - \frac{9\pi}{16}\left(1 - \frac{\pi}{32}\right)\left(\frac{\mu H}{c}\right)^2\right\}. \qquad (6.19б)$$

Поправочный член равен, примерно $1{,}6\left(\frac{\mu H}{c}\right)^2$. Аналогично может быть определен $\varkappa(H)$ для других механизмов рассеяния.

2. Мы определили постоянную Холла R (5.5а), магнетосопротивление $\Delta\rho/\rho$ (5.6а), постоянную Нернста Q (6.4а), изменение термоэдс в магнитном поле $\Delta\alpha$ (6.13) и коэффициент теплопроводности $\varkappa(H)$ при условии $\nabla_y T = 0$, т. е. постоянства температуры в поперечном направлении. В этих условиях соответствующие коэффициенты получили название *изотермических*. Однако в ряде случаев условия опыта больше соответствуют предположению об отсутствии поперечного теплового потока $w_y = 0$; тогда коэффициенты называются *адиабатическими*.

Для определения адиабатических коэффициентов надо исходить из уравнений (6.3) и уравнения (6.18) для

$$w_y = 0 = \varphi j_y - \frac{en}{m} k_0 T \langle\chi_y^* x\rangle. \qquad (6.20)$$

Например, для определения адиабатической постоянной Холла

$R_{ад}$ надо положить $j_x=j$, $j_y=0$ и $\nabla_x T=0$; затем надо из уравнения (6.20) и уравнения $j_y=0$ (6.3) выразить $\nabla_x\left(\frac{\zeta}{e}-\varphi\right)$ и $\nabla_y T$ через $\nabla_y\left(\frac{\zeta}{e}-\varphi\right)$, подставив их в первое уравнение (6.3) с $j_x=j$, получим из него в линейном приближении по магнитному полю

$$\frac{\nabla_y\left(\frac{\zeta}{e}-\varphi\right)}{Hj}=R_{ад}=R_{из}\left(1+\frac{1}{4}\frac{e}{k_0}\alpha\right), \tag{6.21}$$

где α—дифференциальная термоэдс, а изотермическая постоянная Холла $R_{из}$ равна (5.5а). Если в тепловой поток $\boldsymbol{w}$ включить слагаемое $-\varkappa_0\nabla T$, связанное с теплопроводностью решетки, то формула (6.21), как легко проверить, приобретает вид

$$R_{ад}=R_{из}\left(1+\frac{\frac{1}{4}\frac{e}{k_0}\alpha}{1+\frac{\varkappa_0}{\varkappa_e}}\right), \tag{6.21а}$$

где $\varkappa_e$—коэффициент теплопроводности электронов. Полагая, например, для теллура $\varkappa_e=(2k_0^2T/e)n\mu$, $n=4\cdot10^{16}$ $см^{-3}$, $\mu=10^3$ $см^2/в\cdot сек$, $\varkappa_0=4\cdot10^{-3}$ $кал/град$, убедимся, что $R_{ад}$ на 0,4% больше $R_{из}$, т. е. поправка на адиабатичность мала. Аналогично можно определить адиабатическую постоянную Нернста ($w_y=0$, $\nabla_y T\neq0$); при этом оказывается, что в слабом магнитном поле

$$Q_{ад}=Q_{из}\left(1+\frac{\frac{7}{4}\frac{e}{k_0}\alpha}{1+\frac{\varkappa_0}{\varkappa_e}}\right), \tag{6.22}$$

т. е. поправка в семь раз больше, чем в предыдущем случае.

Таким же образом можно вычислить адиабатические коэффициенты других гальвано- и термомагнитных эффектов как для слабого, так и для сильного магнитного поля.

3. Общие выражения для постоянной Холла R (5.5а), магнетосопротивления $\Delta\rho/\rho$ (5.6а), постоянной Нернста Q (6.4а) и изменения термоэдс в магнитном поле α_H (6.11) справедливы для произвольного магнитного поля. Таким образом, для произвольного магнитного поля вопрос сводится к вычислению коэффициентов a_1, a_2, b_1 и b_2 по общим формулам (5.4а), (5.4б) и (6.3а), (6.3б). При этом определение, например, коэффициентов a_1 и a_2, согласно (5.4а), (5.4б), в случае рассеяния электронов на акустических колебаниях (3.5а) сводится к вычислению интегралов

$$\int_0^\infty\frac{x^2e^{-x}\,dx}{h+x}, \qquad \int_0^\infty\frac{x^{3/2}e^{-x}\,dx}{h+x}, \tag{6.23}$$

зависящих от параметра

$$h=\frac{\tau_A^2}{(k_0T)^3}\left(\frac{eH}{mc}\right)^2, \tag{6.23а}$$

который является функцией магнитного поля H и температуры T.

Интегралы (6.23) могут быть выражены через трансцендентные функции от h: интеграл вероятности и интегральную показательную функцию. Нетрудно также численно построить график (или таблицу) зависимости интегралов от h. Аналогичные интегралы возникают при определении коэффициентов b_1 и b_2, а также в случае других механизмов рассеяния[1]). Если нам известна зависимость коэффициентов a_i и b_i от H и T, то тем самым определена зависимость от магнитного поля и температуры величин R, $\Delta\rho/\rho$, Q, α_H и $\varkappa(H)$.

4. Для полупроводников со смешанной проводимостью рассмотрим только эффект Нернста в слабом магнитном поле.

В линейном приближении по магнитному полю H коэффициенты a_1, a_2, b_1, b_2, как следует из (5.8) и (6.5) равны:

$$a_1=\frac{ne^2}{m_n}\langle\tau_n\rangle+\frac{pe^2}{m_p}\langle\tau_p\rangle, \tag{6.24а}$$

$$a_2=\frac{ne^2}{m_n}(\gamma_nH)\langle\tau_n^2\rangle-\frac{pe^2}{m_p}\langle\gamma_pH\rangle\langle\tau_p^2\rangle, \tag{6.24б}$$

$$b_1=\frac{ne^2}{m_n}\left[\left(\frac{k_0}{e}\right)\langle\tau_nx\rangle-\frac{\zeta}{eT}\langle\tau_n\rangle\right]-\frac{pe^2}{m_p}\left[\left(\frac{k_0}{e}\right)\langle\tau_px\rangle+\frac{\varepsilon_G+\zeta}{eT}\langle\tau_p\rangle\right], \tag{6.24в}$$

$$b_2=\frac{ne^2}{m_n}(\gamma_nH)\left[\left(\frac{k_0}{e}\right)\langle\tau_n^2x\rangle-\frac{\zeta}{eT}\langle\tau_n^2\rangle\right]+ +\frac{pe^2}{m_p}(\gamma_pH)\left[\left(\frac{k_0}{e}\right)\langle\tau_p^2x\rangle+\frac{\varepsilon_G+\zeta}{eT}\langle\tau_p^2\rangle\right]. \tag{6.24г}$$

В том же приближении по магнитному полю постоянная Нернста

$$\begin{aligned}Q=\frac{a_1b_1-a_2b_1}{(a_1^2+a_2^2)H}=&\left[\frac{ne^2}{m_n}\langle\tau_n\rangle+\frac{pe^2}{m_p}\langle\tau_p\rangle\right]^{-2}\times\\ &\times\left\{\left(\frac{ne^2}{m_n}\right)^2\gamma_n\left(\frac{k_0}{e}\right)[\langle\tau_n^2x\rangle\langle\tau_n\rangle-\langle\tau_nx\rangle\langle\tau_n^2\rangle]+\right.\\ &+\left(\frac{pe^2}{m_p}\right)^2\gamma_p\left(\frac{k_0}{e}\right)[\langle\tau_p^2x\rangle\langle\tau_p\rangle-\langle\tau_px\rangle\langle\tau_p^2\rangle]+\\ &+\left(\frac{ne^2}{m_n}\right)\left(\frac{pe^2}{m_p}\right)\left[\gamma_p\left(\frac{k_0}{e}\right)(\langle\tau_p^2x\rangle\langle\tau_n\rangle+\langle\tau_p^2\rangle\langle\tau_nx\rangle)+\right.\\ &+\gamma_n\left(\frac{k_0}{e}\right)(\langle\tau_n^2x\rangle\langle\tau_p\rangle+\langle\tau_n^2\rangle\langle\tau_px\rangle)+\\ &\left.\left.+\frac{\varepsilon_G}{eT}(\gamma_n\langle\tau_n^2\rangle\langle\tau_p\rangle+\gamma_p\langle\tau_p^2\rangle\langle\tau_n\rangle)\right]\right\}.\end{aligned} \tag{6.25}$$

[1]) Johnson V. A., Whitesell W. J.—Phys. Rev. 1953, v. 89 p. 941.

Фигурная скобка в этом выражении состоит из трех слагаемых, пропорциональных n^2, p^2 и np. Полагая концентрацию дырок $p=0$, получим из (6.25) выражение (6.6).

Если τ_n, τ_p зависят от энергии по степенному закону (3.4), то, воспользовавшись (3.4а), можно все угловые скобки в (6.25) выразить через Г-функции. Далее, воспользовавшись (3.3), можно вместо e/m_n и e/m_p ввести соответствующие подвижности μ_n и μ_p. Для того чтобы не загромождать изложение, мы приведем результат только для рассеяния электронов и дырок на акустических колебаниях:

$$Q=-\frac{3\pi}{8ce}\left[\frac{k_0}{2}\frac{n\mu_n^3+p\mu_p^3}{(n\mu_n+p\mu_p)^2}-\frac{\varepsilon_G+{}^7/_2k_0T}{T}\frac{n\mu_n\,p\mu_p(\mu_n+\mu_p)}{(n\mu_n+p\mu_p)^2}\right]. \quad (6.25а)$$

§ 7. Явления переноса в полупроводниках с простой зоной при произвольном вырождении

1. Обобщим некоторые результаты, полученные в §§ 3—6 на случай произвольного вырождения носителей тока. Заменим в выражении для плотности тока электронов (2.15) u_n на (2.13а), χ_n на (2.16) и перейдем в интеграле посредством (2.13) от переменной k к ε, тогда [1])

$$\boldsymbol{j}=\frac{2\sqrt{2}}{3\pi^2}\frac{m^{1/2}e^2}{\hbar^3}\int_0^\infty \chi^*(\varepsilon)\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\varepsilon^{3/2}\,d\varepsilon. \quad (7.1)$$

Запишем это выражение, аналогично (2.21), в виде

$$\boldsymbol{j}=\frac{ne^2}{m}\langle\chi^*\rangle, \quad (7.1а)$$

где символ усреднения $\langle\ldots\rangle$ имеет теперь следующий смысл:

$$\langle\chi^*\rangle=\frac{1}{\zeta_0^{3/2}}\int_0^\infty \chi^*(\varepsilon)\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\varepsilon^{3/2}\,d\varepsilon=\frac{1}{z_0^{3/2}}\int_0^\infty \chi^*(x)\left(-\frac{\partial f_0}{\partial x}\right)x^{3/2}\,dx. \quad (7.1б)$$

Здесь ζ_0 — граничная энергия Ферми (VI.2.8а), $z_0=\zeta_0/k_0T$ и $x=\varepsilon/k_0T$. При произвольном вырождении равновесная функция распределения

$$f_0=\frac{1}{\exp\dfrac{\varepsilon-\zeta}{k_0T}+1}=\frac{1}{e^{x-z}+1}, \quad (7.1в)$$

где $z=\zeta/k_0T$; в этом смысле (7.1б) является обобщением выражения (2.19), в которое (7.1б) переходит, если заменить f_0 распределением Больцмана (VI.2.11а).

[1]) Во всех выражениях этого параграфа мы опускаем индекс n.

Из (7.1б) для $\boldsymbol{\chi}^* = \boldsymbol{a} = \text{const}$ следует

$$\langle \boldsymbol{a} \rangle = \boldsymbol{a}. \tag{7.1г}$$

Это может быть просто показано, если взять интеграл (7.1б) по частям и воспользоваться условием нормировки (VI.2.5). Если τ зависит от энергии ε по закону (3.4) и рассматриваются только предельные случаи слабого и сильного магнитного поля, то усреднение в равенстве (7.1а) всегда сводится к усреднению степеней энергии, т. е.

$$\langle \varepsilon^s \rangle = \zeta_0^{-3/2} \int_0^\infty \varepsilon^{s+3/2} \left(-\frac{\partial f_0}{\partial \varepsilon} \right) d\varepsilon, \tag{7.2}$$

$$\langle x^s \rangle = z_0^{-3/2} \int_0^\infty x^{s+3/2} \left(-\frac{\partial f_0}{\partial x} \right) dx. \tag{7.2а}$$

Если $k_0 T/\zeta_0 \ll 1$, то в первом (δ-функционном) и втором приближениях по вырождению (7.2) соответственно равно

$$\langle \varepsilon^s \rangle = \zeta_0^s, \tag{7.2б}$$

$$\langle \varepsilon^s \rangle = \zeta_0^s \left\{ 1 + \frac{\pi^2}{6} s \left(s + \frac{3}{2} \right) \left(\frac{k_0 T}{\zeta_0} \right)^2 \right\}, \tag{7.2в}$$

где мы воспользовались формулой (П.17.5) и выражением для ζ (VI.2.9).

Интегрируя по частям в (7.2а), получим

$$\langle x^s \rangle = z_0^{-3/2} (s + {}^3/_2) \mathscr{F}_{s+1/2}(z), \tag{7.2г}$$

где $z = \zeta/k_0 T$ и интеграл Ферми $\mathscr{F}_{s+1/2}$ определяется равенством (VI.2.6).

Плотность потока энергии может быть получена из (7.1), если заменить (под знаком интеграла) заряд электрона $-e$ на его полную энергию $\varepsilon - e\varphi$, где φ — потенциал электрического поля, действующего на электроны. Таким образом, плотность потока энергии

$$\boldsymbol{w} = -\frac{ne}{m} \langle \boldsymbol{\chi}^* (\varepsilon - e\varphi) \rangle = -\frac{ne}{m} \langle \boldsymbol{\chi}^* \varepsilon \rangle + \varphi \boldsymbol{j}. \tag{7.3}$$

Если в условиях измерения электрический ток равен нулю, то $\boldsymbol{w}$ равно первому слагаемому правой части (7.3).

2. Электрический ток в отсутствие градиента температуры и магнитного поля определяется из равенства (7.1а), где $\boldsymbol{\chi}^*$, согласно (3.1), равно $\tau \boldsymbol{E}$; таким образом,

$$\boldsymbol{j} = \frac{ne^2}{m^*} \langle \tau \rangle \boldsymbol{E}, \tag{7.4}$$

откуда электропроводность

$$\sigma = \frac{ne^2}{m} \langle \tau \rangle. \tag{7.5}$$

Если τ равно (3.4), то

$$\langle\tau\rangle = a\langle\varepsilon^r\rangle = a(k_0T)^r\langle x^r\rangle. \tag{7.6}$$

Используя (7.2г), получим при произвольном вырождении

$$\sigma = \frac{2\sqrt{2}}{3\pi^2}\frac{e^2m^{1/2}}{\hbar^3}a(k_0T)^{r+3/2}\left(r+\frac{3}{2}\right)\mathcal{F}_{r+1/2}(z), \tag{7.7}$$

где $z=\zeta/k_0T$ при $n=\text{const}$ определяется из (VI.2.5), а при наличии доноров и акцепторов — из (VI.2.15).

При сильном вырождении достаточно использовать первое приближение (7.2б); тогда

$$\sigma = \frac{ne^2}{m}\tau(\zeta_0), \tag{7.8}$$

где $\tau(\zeta_0)=\tau(\varepsilon)\big|_{\varepsilon=\zeta_0}$.

Из выражения (7.1) видно, что из-за наличия под интегралом множителя $\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\approx\delta(\varepsilon-\zeta_0)$ в токе принимают участие только электроны, расположенные в слое толщиной порядка k_0T вблизи поверхности Ферми $\varepsilon=\zeta_0$. В самом деле, для других электронов $\delta(\varepsilon-\zeta_0)\approx 0$. В силу этого при вырождении не вводят понятия подвижности электронов $\mu=\sigma/ne$.

3. Выражение для дифференциальной термоэдс (4.5) может быть записано в виде

$$\alpha=\left(\frac{k_0}{e}\right)\left\{\frac{\langle\tau x\rangle}{\langle\tau\rangle}-z\right\}=\left(\frac{k_0}{e}\right)\frac{1}{k_0T}\left\{\frac{\langle\tau\varepsilon\rangle}{\langle\tau\rangle}-\zeta\right\}. \tag{7.9}$$

Если τ равно (3.4), то при произвольном вырождении

$$\alpha=\left(\frac{k_0}{e}\right)\left\{\frac{(r+5/2)\mathcal{F}_{r+3/2}(z)}{(r+3/2)\mathcal{F}_{r+1/2}(z)}-z\right\}. \tag{7.10}$$

Из равенства (7.9) непосредственно видно, что в первом приближении по вырождению $\alpha=0$. В самом деле, первое слагаемое в фигурной скобке (7.9), в δ-функционном приближении, равно

$$\frac{\langle\tau\varepsilon\rangle}{\langle\tau\rangle}=\frac{\tau(\zeta_0)\zeta_0}{\tau(\zeta_0)}=\zeta_0$$

и, следовательно, сокращается со вторым слагаемым. Во втором приближении по вырождению, используя (7.2г) и (VI.2.9), получим [1])

$$\alpha=\frac{\pi^2}{3}\left(r+\frac{3}{2}\right)\left(\frac{k_0}{e}\right)\left(\frac{k_0T}{\zeta_0}\right). \tag{7.11}$$

---

[1]) При этом здесь и дальше мы используем разложение ($x\ll 1$)

$$\frac{1+ax^2}{1+bx^2}\approx(1+ax^2)(1-bx^2)\approx 1+(a-b)x^2$$

с точностью до величин порядка x^2.

Для различных механизмов рассеяния r имеет разные значения (3.5). Для хороших металлов, как это следует из (VIII.4.14), $r=3/2$.

Из (7.11) сразу видно, что для металлов α существенно меньше, чем для невырожденных полупроводников, так как при сильном вырождении α имеет лишний множитель (k_0T/ζ_0), который при комнатной температуре для металлов порядка $5\cdot10^{-3}$.

4. Подставляя в (4.20) (3.3), (4.3а) и (4.11а), получим для коэффициента теплопроводности

$$\varkappa=\frac{n}{m}k_0^2T\frac{\langle\tau\rangle\langle\tau x^2\rangle-\langle\tau x\rangle^2}{\langle\tau\rangle}=\frac{n}{m}\frac{1}{T}\frac{\langle\tau\rangle\langle\tau\varepsilon^2\rangle-\langle\tau\varepsilon\rangle^2}{\langle\tau\rangle}. \tag{7.12}$$

Легко показать, аналогично тому, как это было сделано для α, что $\varkappa$ в первом приближении по вырождению равно нулю. Во втором приближении по вырождению при τ, равном (3.4),

$$\varkappa=\frac{\pi^2}{3}k_0^2T\frac{n}{m}\tau(\zeta_0), \tag{7.13}$$

где было использовано разложение (7.2в). Отсюда и из (7.8) следует закон Видемана—Франца:

$$\varkappa/\sigma=LT, \tag{7.14}$$

где в случае сильного вырождения число Лорентца

$$L=\frac{\pi^2}{3}\left(\frac{k_0}{e}\right)^2 \tag{7.14а}$$

не зависит от механизма рассеяния.

При промежуточном вырождении $\varkappa$ может быть посредством (7.2г) выражено через интегралы Ферми. В общем случае число Лорентца

$$L=\frac{(r+3/2)(r+7/2)\mathcal{F}_{r+1/2}\mathcal{F}_{r+5/2}-(r+5/2)^2\mathcal{F}^2_{r+3/2}}{(r+3/2)^2\mathcal{F}^2_{r+3/2}}\left(\frac{k_0}{e}\right)^2. \tag{7.15}$$

5. Рассмотрим основные гальваномагнитные явления. Если магнитное поле перпендикулярно вектору $\boldsymbol{P}=\tau\boldsymbol{E}$, то имеют место уравнения (5.4), (5.4а), (5.4б). Конечно, мы должны при этом под символом $\langle\,\rangle$ понимать усреднение (7.1б). Таким образом, формулы для постоянной Холла R и магнетосопротивления $\Delta\rho/\rho$, полученные в § 5 для слабых и сильных магнитных полей, остаются в силе, если только под $\langle\,\rangle$ понимать усреднение (7.1б).

А. Слабые поля: $\gamma\tau H\ll1$. В этом случае, согласно выражению (5.10),

$$R=-\frac{\langle\tau^2\rangle}{\langle\tau\rangle^2}\frac{1}{ecn}. \tag{7.16}$$

В первом приближении по вырождению

$$\frac{\langle\tau^2\rangle}{\langle\tau\rangle^2}=\frac{\tau^2(\zeta_0)}{[\tau(\zeta_0)]^2}=1,$$

поэтому

$$R = -1/ecn. \tag{7.17}$$

В случае зависимости (3.4) и промежуточного вырождения множитель $\langle\tau^2\rangle/\langle\tau\rangle^2$ может быть выражен через интегралы Ферми.

Магнетосопротивление, согласно (5.11), равно

$$\frac{\Delta\rho}{\rho} = \left(\frac{eH}{mc}\right)^2 \frac{\langle\tau^3\rangle\langle\tau\rangle - \langle\tau^2\rangle^2}{\langle\tau\rangle^2}, \tag{7.18}$$

что в первом приближении по вырождению равно нулю. Это соответствует тому общему положению, что магнетосопротивление возникает только тогда, когда имеются группы электронов разной энергии, которые по-различному рассеиваются.

Во втором приближении по вырождению при степенной зависимости времени релаксации от энергии (3.4) получим

$$\Delta\rho/\rho = B_0 H^2, \tag{7.19}$$

где

$$B_0 = \frac{\pi^2}{3} r^2 \left[\frac{e\tau(\zeta_0)}{mc}\right]^2 \left(\frac{k_0 T}{\zeta_0}\right)^2. \tag{7.19а}$$

В случае промежуточного вырождения величина $\Delta\rho/\rho$ (7.18) может быть выражена через интегралы Ферми.

Отметим, что в металлах при слабом поле и сильном вырождении $\Delta\rho/\rho$ от температуры не зависит, так как $\tau(\zeta_0)$ обратно пропорционально температуре (VIII.4.14).

Б. Сильные поля: $\gamma\tau H \gg 1$. Для сильных полей независимо от степени вырождения постоянная Холла R равна (7.17), как это следует из (5.16). Согласно равенству (5.17) магнетосопротивление

$$\frac{\Delta\rho}{\rho} = \langle\tau\rangle\left\langle\frac{1}{\tau}\right\rangle - 1. \tag{7.20}$$

В первом приближении по вырождению $\Delta\rho/\rho = 0$, так как $\langle\tau\rangle\left\langle\frac{1}{\tau}\right\rangle = 1$. В случае (3.4) во втором приближении по вырождению получим

$$\frac{\Delta\rho}{\rho} = \frac{\pi^2}{3} r^2 \left(\frac{k_0 T}{\zeta_0}\right)^2, \tag{7.21}$$

т. е. насыщение при $H \to \infty$. При сильных полях и вырождении величина насыщения $\Delta\rho/\rho \sim T^2$.

В. Произвольные поля: случай сильного вырождения. $\Delta\rho/\rho$ может быть вычислено для произвольного магнитного поля. Во втором приближении по вырождению

$$\frac{\Delta\rho}{\rho} = \frac{B_0 H^2}{1 + (R\sigma H)^2}. \tag{7.22}$$

Здесь B_0 определено формулой (7.19а), а R и σ—постоянная Холла (7.17) и электропроводность (7.8). Из (7.22) для больших

магнитных полей

$$\left(\frac{\Delta\rho}{\rho}\right)_{H\to\infty} = \frac{B_0}{(R\sigma)^2}, \tag{7.22a}$$

что совпадает с (7.21).

Экспериментальные данные о магнетосопротивлении металлов плохо согласуются с формулой (7.22). При самых малых полях $\Delta\rho/\rho$ действительно пропорционально H^2, но далее возрастание $\Delta\rho/\rho$ становится пропорциональным H и к настоящему насыщению не стремится. Кроме того, наблюдаемые на опыте B_0 на несколько порядков больше теоретических. Температурная зависимость $\Delta\rho/\rho$, предсказываемая теорией, на опыте не оправдывается. Главная причина всех этих расхождений заключается, по-видимому, в сложной структуре энергетического спектра электронов.

И. М. Лифшиц с сотрудниками[1]) развил теорию гальваномагнитных явлений в металлах, учитывая сложную форму поверхности Ферми ($\varepsilon = \zeta_0$). Ими было показано, что асимптотическое поведение компонент гальваномагнитного тензора в сильных магнитных полях существенно связано с топологией поверхности Ферми и в первую очередь с тем, является ли поверхность Ферми открытой или закрытой, т. е. проходит ли она непрерывно через все $\boldsymbol{k}$-пространство или распадается на замкнутые поверхности, повторяющиеся в пространстве обратной решетки.

6. Из термомагнитных явлений мы рассмотрим только поперечный эффект Нернста. Так же, как в случае гальваномагнитных явлений, общая структура уравнений (6.3) и выражения для коэффициентов b_1 и b_2 (6.3а), (6.3б) сохраняются при любой степени вырождения носителей тока, если только под символом $\langle\ \rangle$ понимать усреднение (7.1б). Это означает, что для постоянной Нернста Q мы можем воспользоваться выражением (6.4а) и следующими из нее выражениями для слабых и сильных магнитных полей.

А. Слабые поля: $\gamma\tau H \ll 1$. Из (6.6а) следует, что

$$Q = \left(\frac{k_0}{e}\right)\frac{e}{mc}\,\frac{\langle\tau\rangle\langle\tau^2 x\rangle - \langle\tau^2\rangle\langle\tau x\rangle}{\langle\tau\rangle^2}. \tag{7.23}$$

Отсюда непосредственно видно, что в первом приближении по вырождению $Q = 0$. Это отражает то общее положение, что для того, чтобы имел место эффект Нернста, необходимо наличие различных групп электронов, которые по-разному рассеиваются. Если τ от ε не зависит, то Q также равно нулю. В случае (3.4)

---

[1]) Лифшиц И. М., Азбель М. Я., Каганов М. И. Электронная теория металлов.— М.: Наука, 1971.

во втором приближении по вырождению

$$Q=\frac{\pi^2}{3}r\left(\frac{k_0}{e}\right)\left(\frac{e\tau(\zeta_0)}{mc}\right)\left(\frac{k_0T}{\zeta_0}\right), \tag{7.24}$$

где было использовано разложение (7.2в).

Отсюда, в частности, видно, что знак эффекта, так же как и в невырожденном случае, совпадает со знаком r. Наличие множителя k_0T/ζ_0 обусловливает малость эффекта. В промежуточной области вырождения Q может быть выражено через интегралы Ферми.

Б. Сильные поля: $\gamma\tau H\gg 1$. Из выражения (6.9) видно, что в первом приближении по вырождению $Q=0$. Во втором приближении по вырождению при степенной зависимости времени релаксации от энергии (3.4) получим

$$Q=\frac{\pi^2}{3}r\left(\frac{k_0}{e}\right)\frac{mc}{e\tau(\zeta_0)H^2}\left(\frac{k_0T}{\zeta_0}\right). \tag{7.25}$$

По поводу этого выражения можно сказать все то, что было сказано о (7.24).

Изменение термоэдс α в магнитном поле может быть рассмотрено совершенно аналогично предыдущему случаю.

§ 8. Явления переноса в полупроводниках типа германия и кремния

1. Рассмотрим сначала явления переноса в электронном полупроводнике в случае, когда минимум энергии не расположен в центре бриллюэновской зоны. В этом случае поверхности постоянной энергии будут, вообще говоря, не сферами, а эллипсоидами. При этом вблизи минимума энергия электрона $\varepsilon(k_x, k_y, k_z)$ зависит от составляющих волнового вектора, как и в случае простой зоны, квадратично.

Как мы видели в гл. IV, § 15 зона проводимости кремния имеет шесть симметрично расположенных минимумов энергии в точках на осях Δ, т. е. в направлениях [100]; зона проводимости германия обладает 8 минимумами энергии, расположенными на границах зоны Бриллюэна в направлениях [111]. Изоэнергетические поверхности электрона $\varepsilon(\boldsymbol{k})=\text{const}$ вблизи этих минимумов имеют вид эллипсоидов вращения (IV.15.1). Такие энергетические зоны получили название *многоэллипсоидных* или *многодолинных*. Затруднения, возникающие при наличии многоэллипсоидной зонной структуры, связаны не только с усложнением выражений для кинетических коэффициентов, но и с тем, что время релаксации τ, если оно вообще может быть введено, перестает быть скаляром и превращается в тензор.

Однако, по-видимому, во многих случаях можно и при сложной зонной структуре считать время релаксации электрона скаляром,

зависящим от его энергии. В этом параграфе мы ограничимся этим приближением.

Так как расстояние между эквивалентными минимумами в $\boldsymbol{k}$-пространстве порядка постоянной обратной решетки, то электрон при поглощении фонона может перейти из области одного минимума в область другого только в том случае, если он поглотит фонон с волновым вектором $q \approx 1/a$, где a — постоянная решетки (закон сохранения волнового вектора).

При низких температурах такие межэллипсоидные переходы будут относительно редки, но при более высоких температурах они могут, как показал Херринг, в некоторых случаях объяснить более сильную зависимость подвижности от температуры ($\mu \sim T^{-n}$, где $n > 1{,}5$). Теория межэллипсоидных переходов разработана еще довольно слабо, поэтому мы в дальнейшем этот тип рассеяния рассматривать не будем. В случае же внутриэллипсоидного рассеяния электронов токи от всех эквивалентных эллипсоидов складываются независимо.

2. Если отсчитывать энергию электрона ε от значения, соответствующего одному из эквивалентных минимумов, а волновой вектор электрона $\boldsymbol{k}$ от точки i-го минимума в $\boldsymbol{k}$-пространстве, то в главных осях i-го эллипсоида энергии

$$\varepsilon = \frac{\hbar^2}{2}\left[\frac{k_1^2}{m_1} + \frac{k_2^2}{m_2} + \frac{k_3^2}{m_3}\right] = \frac{\hbar^2}{2}\sum_{\alpha=1}^{3}\frac{k_\alpha^2}{m_\alpha}, \tag{8.1}$$

где k_α — прямоугольные составляющие волнового вектора $\boldsymbol{k}$, а m_α — компоненты тензора эффективной массы (IV.15.1). В случае германия и кремния, направляя ось z ($\alpha = 3$) вдоль оси вращения эллипсоида энергии и полагая $m_3 = m_\parallel$ — продольной массе, $m_1 = m_2 = m_\perp$ — поперечной массе, получим

$$\varepsilon = \frac{\hbar^2}{2}\left[\frac{k_1^2 + k_2^2}{m_\perp} + \frac{k_3^2}{m_\parallel}\right]. \tag{8.1a}$$

Для дальнейшего удобно ввести переменные

$$w_\alpha = \frac{\hbar}{\sqrt{2m_\alpha}}\, k_\alpha, \tag{8.2}$$

в которых

$$\varepsilon = w_1^2 + w_2^2 + w_3^2. \tag{8.2a}$$

Составляющие скорости электрона равны

$$v_\alpha = \frac{1}{\hbar}\frac{\partial \varepsilon}{\partial k_\alpha} = \frac{\hbar}{m_\alpha}\, k_\alpha = \sqrt{\frac{2}{m_\alpha}}\, w_\alpha. \tag{8.3}$$

Кинетическое уравнение (VIII.2.6) остается справедливым и в случае сложной зонной структуры, поэтому

$$\boldsymbol{v}\nabla_{\boldsymbol{r}} f - \frac{e}{\hbar}\left\{\boldsymbol{E} + \frac{1}{c}[\boldsymbol{v}\boldsymbol{H}]\right\}\nabla_{\boldsymbol{k}} f = -\frac{f - f_0}{\tau(\varepsilon)}. \tag{8.4}$$

Здесь мы, как указано выше, предполагаем, что существует время релаксации, зависящее от энергии, так что член столкновений

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -\frac{f-f_0}{\tau(\varepsilon)}. \tag{8.4a}$$

3. Рассмотрим электропроводность, обусловленную одним эллипсоидом энергии при наличии одного электрического поля ($\nabla T = 0$, $\boldsymbol{H} = 0$). В этом случае кинетическое уравнение (8.4) имеет вид

$$\frac{e\tau}{\hbar}\boldsymbol{E}\nabla_{\boldsymbol{k}} f = f - f_0. \tag{8.5}$$

Положим неравновесную функцию распределения

$$f = f_0 + \boldsymbol{f}^{(10)}\boldsymbol{E} = f_0 + \sum_{\mu=1}^{3} f_\mu^{(10)} E_\mu, \tag{8.6}$$

что можно рассматривать как разложение f в ряд по степеням электрического поля E_μ с удержанием членов первой степени; так как неравновесная добавка к функции распределения — скалярная величина, то $\boldsymbol{f}^{(10)}$ — вектор, т. е. $f_\mu^{(10)}$ — компоненты тензора 1-го ранга (Приложение 11). Подставляя (8.6) в (8.5) и отбрасывая члены порядка E^2, получим

$$\frac{e\tau}{\hbar}\boldsymbol{E}\nabla_{\boldsymbol{k}} f_0 = \boldsymbol{E}\boldsymbol{f}^{(10)}, \tag{8.7}$$

откуда (ввиду произвольности $\boldsymbol{E}$)

$$\boldsymbol{f}^{(10)} = \frac{e\tau}{\hbar}\nabla_{\boldsymbol{k}} f_0. \tag{8.8}$$

Так как равновесная функция распределения f_0 зависит от $\boldsymbol{k}$ только через энергию $\varepsilon(\boldsymbol{k})$, то

$$\nabla_{\boldsymbol{k}} f_0 = \frac{\partial f_0}{\partial \varepsilon}\nabla_{\boldsymbol{k}}\varepsilon = \frac{\partial f_0}{\partial \varepsilon}\hbar\boldsymbol{v}, \tag{8.9}$$

так что

$$\boldsymbol{f}^{(10)} = e\tau(\varepsilon)\frac{\partial f_0}{\partial \varepsilon}\boldsymbol{v} \tag{8.10}$$

или

$$f_\mu^{(10)} = e\tau(\varepsilon)\frac{\partial f_0}{\partial \varepsilon}v_\mu. \tag{8.10a}$$

Электрический ток, создаваемый электронами i-го эллипсоида,

$$\boldsymbol{j}^{(i)} = -e\sum_{(\boldsymbol{k})} f\boldsymbol{v} = e^2\sum_{(\boldsymbol{k})}(\boldsymbol{E}\boldsymbol{v})\boldsymbol{v}\tau(\varepsilon)\left(-\frac{\partial f_0}{\partial \varepsilon}\right), \tag{8.11}$$

так как равновесная функция распределения f_0 тока не дает. Здесь

$$\sum_{(\boldsymbol{k})} \equiv \frac{1}{4\pi^3}\int d^3k \tag{8.11a}$$

обозначает суммирование (интегрирование) по волновому вектору $\boldsymbol{k}$. Составляющая тока, согласно (8.11), равна

$$j_\alpha^{(i)} = e^2 \sum_{(k)} \sum_{\beta} \tau(\varepsilon) \left(-\frac{\partial f_0}{\partial \varepsilon}\right) v_\alpha v_\beta E_\beta = \sum_{\beta} \sigma_{\alpha\beta}^{(i)} E_\beta, \tag{8.12}$$

где

$$\sigma_{\alpha\beta}^{(i)} = e^2 \sum_{(k)} \tau(\varepsilon) \left(-\frac{\partial f_0}{\partial \varepsilon}\right) v_\alpha v_\beta \tag{8.12а}$$

— компонента тензора электропроводности i-го эллипсоида. Так как $v_\alpha \infty k_\alpha$ и ε — четная функция k_α, то $\sigma_{\alpha\beta} = 0$, если $\alpha \neq \beta$, т. е. в главных осях эллипсоида энергии тензор электропроводности диагонален. Таким образом,

$$\sigma_{\alpha\alpha}^{(i)} = e^2 \sum_{(k)} \tau(\varepsilon) \left(-\frac{\partial f_0}{\partial \varepsilon}\right) v_\alpha^2 = e^2 \sum_{(w)} \tau(\varepsilon) \left(-\frac{\partial f_0}{\partial \varepsilon}\right) \frac{2}{m_\alpha} w_\alpha^2, \tag{8.12б}$$

где мы перешли к суммированию по $\boldsymbol{w}$. Учитывая изотропную связь между ε и w_α (8.2а), легко показать, что суммирование по $\boldsymbol{w}$ сводится к суммированию по ε, с заменой w_α^2 на $^1/_3\varepsilon$. Тогда

$$\sigma_{\alpha\alpha}^{(i)} = \frac{e^2}{m_\alpha} \frac{2}{3} \sum_{(\varepsilon)} \tau(\varepsilon) \left(-\frac{\partial f_0}{\partial \varepsilon}\right) \varepsilon. \tag{8.12в}$$

Используя выражение для плотности состояний (VI.2.22), получим

$$\frac{2}{3} \sum_{(\varepsilon)} \tau(\varepsilon) \left(-\frac{\partial f_0}{\partial \varepsilon}\right) \varepsilon =$$

$$= \frac{2\sqrt{2}}{3\pi^2} \frac{(m_1 m_2 m_3)^{1/2}}{\hbar^3} \int_0^\infty \tau(\varepsilon) \left(-\frac{\partial f_0}{\partial \varepsilon}\right) \varepsilon^{3/2}\, d\varepsilon = n^{(i)} \langle \tau \rangle, \tag{8.13}$$

где $n^{(i)}$ — число электронов в одном эллипсоиде, а символ $\langle\ \rangle$ прямое обобщение выражения (7.1б), в которое он переходит при $m_1 = m_2 = m_3 = m$. Таким образом,

$$\sigma_{\alpha\alpha}^{(i)} = \frac{e^2 n^{(i)}}{m_\alpha} \langle \tau \rangle, \tag{8.14}$$

что сходно с (7.5).

Выберем теперь прямоугольную координатную систему, оси которой направлены по ребрам куба элементарной ячейки кристалла. Компоненты тензора полной электропроводности (от всех эллипсоидов) в этой координатной системе

$$\sigma_{\lambda\mu} = \sum_i \sigma_{\lambda\mu}^{(i)}, \tag{8.15}$$

где $\sigma_{\lambda\mu}^{(i)}$ — компоненты тензора (8.14), преобразованные к координатной системе, связанной с осями кристалла.

В кубическом кристалле тензор электропроводности $\sigma_{\lambda\mu}$ равен скаляру σ.

Рассмотрим вначале простой случай кремния, у которого минимумы энергии расположены по направлениям $\langle 100\rangle$. Этим же направлениям параллельны оси вращения эллипсоидов энергии (эффективные массы $m_{\parallel}$).

Таким образом, в случае кремния главные оси всех шести эллипсоидов энергии параллельны осям прямоугольной координатной системы, совпадающим с ребрами куба элементарной ячейки. Из (8.15) и (8.14) следует

$$\sigma = \sigma_{xx} = 2\frac{e^2 n^{(i)}}{m_{\parallel}}\langle\tau\rangle + 4\frac{e^2 n^{(i)}}{m_{\perp}}\langle\tau\rangle = \frac{e^2 n}{m'}\langle\tau\rangle, \tag{8.16}$$

где концентрация электронов

$$n = N_c n^{(i)} = 6n^{(i)} \tag{8.16а}$$

(N_c — число эквивалентных минимумов) и

$$\frac{1}{m'} = \frac{1}{3}\left(\frac{1}{m_{\parallel}} + \frac{2}{m_{\perp}}\right). \tag{8.16б}$$

Пользуясь формулами преобразования компонент тензора (Приложение 11) можно показать, что формулы (8.16а), (8.16б) справедливы и в случае германия, когда главные оси эллипсоидов энергии не параллельны ребрам куба.

Отметим, что выражение (8.16б) следует и из того, что $1/m'$ должно быть пропорционально скалярному инварианту тензора обратной эффективной массы, т. е. его шпуру (следу) $\sum_{\alpha} 1/m_{\alpha}$ (множитель 1/3 может быть обоснован рассмотрением случая скалярной эффективной массы).

Подвижность

$$\mu = \frac{\sigma}{en} = \frac{e\langle\tau\rangle}{m'}, \tag{8.17}$$

что сходно с (3.3).

4. Рассмотрим эффект Холла в слабом магнитном поле. Для этого надо кинетическое уравнение (8.4) (с $\nabla_r f = 0$) решить с точностью до членов порядка H. Рассматривая опять один эллипсоид, положим в его главных осях функцию распределения

$$f = f_0 + \sum_{\mu} f_{\mu}^{(10)} E_{\mu} + \sum_{\mu\nu} f_{\mu\nu}^{(11)} E_{\mu} H_{\nu}, \tag{8.18}$$

где $f_{\mu\nu}^{(11)}$ — тензор 2-го ранга (только в этом случае добавка пропорциональная H есть скаляр). Подставим (8.18) в левую и правую части кинетического уравнения

$$\frac{e\tau}{\hbar}\left\{\boldsymbol{E} + \frac{1}{c}[\boldsymbol{v}\boldsymbol{H}]\right\}\nabla_k\left(f_0 + \sum_{\alpha} f_{\mu}^{(10)} E_{\mu} + \sum_{\mu\nu} f_{\mu\nu}^{(11)} E_{\mu} H_{\nu}\right) = \sum_{\mu} f_{\mu}^{(10)} E_{\mu} + \sum_{\mu\nu} f_{\mu\nu}^{(10)} E_{\mu} H_{\nu}. \tag{8.19}$$

Согласно предыдущему выражение $(e\tau/\hbar)\,\boldsymbol{E}\nabla_k f_0$ сокращается с первой суммой, стоящей справа.

Так как

$$[\boldsymbol{vH}]\,\nabla_k f_0 = [\boldsymbol{vH}]\frac{\partial f_0}{\partial \varepsilon}\,\nabla_k \varepsilon \sim [\boldsymbol{vH}]\,\boldsymbol{v} = 0,$$

то линейный по E и H член слева равен

$$\frac{e\tau}{\hbar c}[\boldsymbol{vH}]\,\nabla_k\left\{\sum_\mu f_\mu^{(10)} E_\mu\right\} = \frac{e\tau}{\hbar c}[\boldsymbol{vH}]\sum_\mu E_\mu \nabla_k\left(e\tau\frac{\partial f_0}{\partial \varepsilon}\,v_\mu\right) =$$

$$= \frac{e^2\tau}{\hbar c}[\boldsymbol{vH}]\sum_\mu E_\mu\left\{v_\mu\nabla_k\left(\tau\frac{\partial f_0}{\partial \varepsilon}\right) + \tau\frac{\partial f_0}{\partial \varepsilon}\,\nabla_k v_\mu\right\}. \qquad (8.20)$$

Так как τ зависит только от ε, то

$$\nabla_k\left(\tau\frac{\partial f_0}{\partial \varepsilon}\right) = \frac{\partial}{\partial \varepsilon}\left(\tau\frac{\partial f_0}{\partial \varepsilon}\right)\nabla_k\varepsilon = \frac{\partial}{\partial \varepsilon}\left(\tau\frac{\partial f_0}{\partial \varepsilon}\right)\hbar\boldsymbol{v}$$

и произведение $[\boldsymbol{vH}]$ на первое слагаемое в фигурных скобках в (8.20) равно нулю. С другой стороны,

$$\nabla_k v_\mu = \frac{\hbar}{m_\mu}\,\boldsymbol{e}_\mu, \qquad (8.21)$$

если воспользоваться (8.3). Здесь $\boldsymbol{e}_\mu$ — единичный вектор по оси μ. Таким образом, в (8.20) фигурирует скалярное произведение

$$[\boldsymbol{vH}]\,\boldsymbol{e}_\mu = [\boldsymbol{vH}]_\mu = \sum_{\alpha\nu}\delta_{\mu\alpha\nu}v_\alpha H_\nu, \qquad (8.22)$$

где $\delta_{\mu\alpha\nu}$ — символ, равный $+1$, когда $\mu\alpha\nu$ — четная перестановка индексов 1, 2, 3; равный -1, когда $\mu\alpha\nu$ — нечетная перестановка 1, 2, 3, и равный нулю во всех случаях, когда среди индексов $\mu\alpha\nu$ имеются одинаковые [1]).

В результате из (8.19) следует

$$\frac{e^2\tau^2}{c}\frac{\partial f_0}{\partial \varepsilon}\sum_{\mu\alpha\nu}\delta_{\mu\alpha\nu}\frac{v_\alpha}{m_\mu}E_\mu H_\nu = \sum_{\mu\nu} f_{\mu\nu}^{(11)} E_\mu H_\nu. \qquad (8.23)$$

Сравнивая коэффициенты при $E_\mu H_\nu$, получим

$$f_{\mu\nu}^{(11)} = \frac{e^2\tau^2}{c}\frac{\partial f_0}{\partial \varepsilon}\frac{1}{m_\mu}\sum_\alpha \delta_{\mu\alpha\nu}v_\alpha = \frac{e^2\tau^2}{c}\left(-\frac{\partial f_0}{\partial \varepsilon}\right)\frac{1}{m_\mu}\sum_\alpha \delta_{\mu\nu\alpha}v_\alpha. \qquad (8.24)$$

Ток, связанный с добавкой (8.23) к функции распределения, равен

$$j_\lambda^{(i)} = -e\sum_{(k)}\left\{v_\lambda\sum_{\mu\nu} f_{\mu\nu}^{(11)} E_\mu H_\nu\right\} = \frac{e^3}{c}\sum_{\mu\nu\alpha}\delta_{\mu\nu\alpha}E_\mu H_\nu\sum_{(k)}\tau^2\frac{\partial f_0}{\partial \varepsilon}\frac{v_\alpha v_\lambda}{m_\mu}. \qquad (8.25)$$

---

[1]) См. Приложение 24. Некоторые авторы обозначают $\delta_{\mu\alpha\nu}$ через $\varepsilon_{\mu\alpha\nu}$.

Аналогично (8.12а) выражение (8.25) отлично от нуля при $\alpha = \lambda$; в этом случае можно интегрирование по $\boldsymbol{k}$ свести к интегрированию по ε подобно (8.12в). В результате

$$j_\lambda^{(i)} = -\frac{e^3}{c} \sum_{\mu\nu} \delta_{\mu\nu\lambda} n^{(i)} \langle \tau^2 \rangle \frac{1}{m_\mu m_\lambda} E_\mu H_\nu = \sum_{\mu\nu} \sigma_{\lambda\mu\nu}^{(i)} E_\mu H_\nu, \tag{8.26}$$

где символ $\langle\ \rangle$ определен в (8.13), а тензор 3-го ранга

$$\sigma_{\lambda\mu\nu}^{(i)} = -\frac{e^3}{c} n^{(i)} \langle \tau^2 \rangle \frac{\delta_{\lambda\mu\nu}}{m_\lambda m_\mu} \tag{8.26а}$$

определяет добавочный ток в направлении λ, пропорциональный $E_\mu H_\nu$ [1]). Без учета межэллипсоидных переходов полный ток в координатной системе, связанной с осями кристалла,

$$j_\lambda = \sum_i j_\lambda^{(i)} = \sum_{\mu\nu} \sigma_{\lambda\mu\nu} E_\mu H_\nu, \tag{8.27}$$

где

$$\sigma_{\lambda\mu\nu} = \sum_i \sigma_{\lambda\mu\nu}^{(i)} \tag{8.27а}$$

и $\sigma_{\lambda\mu\nu}^{(i)}$ — компоненты тензора (8.26а), преобразованные к координатной системе, связанной с кристаллом. В кубическом кристалле циклическая перестановка индексов $\lambda\mu\nu$ не меняет значения $\sigma_{\lambda\mu\nu}$, кроме того, естественно думать, что в кубическом кристалле холловский ток пропорционален $[\boldsymbol{EH}]$. Это значит, что

$$\sigma_{\lambda\mu\nu} = \eta \delta_{\lambda\mu\nu}, \tag{8.28}$$

где η — скаляр, через который выражаются все 27 компонент тензора $\sigma_{\lambda\mu\nu}$. Из (8.28) и (8.27)

$$\boldsymbol{j} = \eta [\boldsymbol{EH}]. \tag{8.29}$$

Составим инвариант (скаляр)

$$\sum_{\lambda\mu\nu} \sigma_{\lambda\mu\nu} \delta_{\lambda\mu\nu} = \sum_{\lambda\mu\nu} \eta \delta_{\lambda\mu\nu}^2 = 6\eta \tag{8.30}$$

и выразим его через $\sigma_{\lambda\mu\nu}^{(i)}$

$$6\eta = \sum_{\lambda\mu\nu} \sum_i \sigma_{\lambda\mu\nu}^{(i)} \delta_{\lambda\mu\nu} = \sum_i \sum_{\lambda\mu\nu} \sigma_{\lambda\mu\nu}^{(i)} \delta_{\lambda\mu\nu}. \tag{8.30а}$$

Так как каждая сумма, относящаяся к i-му эллипсоиду, тоже инвариант, то она может быть вычислена в любой координатной системе, в том числе и в главных осях тензора $\sigma_{\lambda\mu\nu}^{(i)}$.

---

[1]) Конечно, $\sigma_{\lambda\mu\nu}^{(i)}$ не имеет размерности удельной электропроводности.

Используя (8.26а), получим

$$\sum_{\lambda\mu\nu}\sigma^{(i)}_{\lambda\mu\nu}\delta_{\lambda\mu\nu}=-\frac{e^3}{c}n^{(i)}\langle\tau^2\rangle\sum_{\lambda\mu\nu}\frac{\delta^2_{\lambda\mu\nu}}{m_\lambda m_\mu}=$$

$$=-\frac{e^3}{c}n^{(i)}\langle\tau^2\rangle 2\left[\frac{1}{m_1m_2}+\frac{1}{m_1m_3}+\frac{1}{m_2m_3}\right], \quad (8.30б)$$

выражение, одинаковое для всех эллипсоидов.

Из (8.30а), (8.30б)

$$\eta=-\frac{e^3}{c}n\langle\tau^2\rangle\frac{1}{m''^2}, \quad (8.31)$$

где

$$\frac{1}{m''^2}=\frac{1}{3}\left[\frac{1}{m_1m_2}+\frac{1}{m_1m_3}+\frac{1}{m_2m_3}\right], \quad (8.31а)$$

что для Ge и Si равно

$$\frac{1}{m''^2}=\frac{1}{3}\left[\frac{2}{m_\perp m_\parallel}+\frac{1}{m_\perp^2}\right]. \quad (8.31б)$$

Из (8.16) и (8.29) следует, что полный ток

$$\boldsymbol{j}=\sigma\boldsymbol{E}+\eta[\boldsymbol{EH}]. \quad (8.32)$$

Исходя из определения постоянной Холла R (5.5), получим из (8.32) в линейном приближении по магнитному полю

$$R=\frac{E_y}{jH}=\frac{\eta}{\sigma^2}=-\frac{1}{ecn}\frac{\langle\tau^2\rangle}{\langle\tau\rangle^2}\left(\frac{m'}{m''}\right)^2, \quad (8.33)$$

что отличается от (6.10) только множителем $(m'/m'')^2$.

5. Рассмотрим магнетопроводимость в слабых магнитных полях. Для этого разложим неравновесную функцию распределения в ряд по напряженности магнитного поля до членов порядка H^2

$$f=f_0+\sum_\mu f^{(10)}_\mu E_\mu+\sum_{\mu\nu}f^{(11)}_{\mu\nu}E_\mu H_\nu+\sum_{\mu\nu\rho}f^{(12)}_{\mu\nu\rho}E_\mu H_\nu H_\rho, \quad (8.34)$$

где $f^{(12)}_{\mu\nu\rho}$—тензор 3-го ранга.

Учитывая значения $f^{(10)}_\mu$ и $f^{(11)}_{\mu\nu}$ очевидно, что для определения $f^{(12)}_{\mu\nu\rho}$ необходимо в левую часть кинетического уравнения подставить

$$\sum_{\mu\nu}f^{(11)}_{\mu\nu}E_\mu H_\nu.$$

При этом мы получим значение, пропорциональное величине

$$[v\boldsymbol{H}]\nabla_k f^{(11)}_{\mu\nu}=-\frac{e^2}{cm_\mu}[v\boldsymbol{H}]\nabla_k\left(\tau^2\frac{\partial f_0}{\partial\varepsilon}\sum_\alpha\delta_{\mu\nu\alpha}v_\alpha\right). \quad (8.35)$$

Производя преобразования, аналогичные (8.20)—(8.22), получим

из кинетического уравнения

$$-\frac{e^3}{c^2}\tau^3\frac{\partial f_0}{\partial\varepsilon}\sum_{\mu\nu}\sum_{\alpha\beta\rho}\frac{\delta_{\mu\nu\alpha}\delta_{\alpha\beta\rho}}{m_\alpha m_\mu}v_\beta E_\mu H_\nu H_\rho=\sum_{\mu\nu\rho}f^{(12)}_{\mu\nu\rho}E_\mu H_\nu H_\rho, \quad (8.36)$$

откуда

$$f^{(12)}_{\mu\nu\rho}=-\frac{e^3}{c^2}\sum_{\alpha\beta}\delta_{\mu\nu\alpha}\delta_{\alpha\beta\rho}\tau^3\frac{\partial f_0}{\partial\varepsilon}\frac{v_\beta}{m_\alpha m_\mu}. \quad (8.37)$$

Определим теперь ток от i-го эллипсоида, обусловленный этой добавкой к функции распределения:

$$j^{(i)}_\lambda=-e\sum_{(\boldsymbol{k})}v_\lambda\sum_{\mu\nu\rho}f^{(12)}_{\mu\nu\rho}E_\mu H_\nu H_\rho=\sum_{\mu\nu\rho}\sigma^{(i)}_{\lambda\mu\nu\rho}E_\mu H_\nu H_\rho. \quad (8.38)$$

Здесь $\sigma^{(i)}_{\lambda\mu\nu\rho}$—тензор 4-го ранга магнетопроводимости, имеющий $3\times3\times3\times3=81$ компоненту. Подставляя (8.36) в (8.38), получим

$$\sigma^{(i)}_{\lambda\mu\nu\rho}=-\frac{e^4}{c^2}\sum_{(\boldsymbol{k})}\sum_{\alpha\beta}\delta_{\mu\nu\alpha}\delta_{\alpha\beta\rho}\tau^3\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\frac{v_\lambda v_\beta}{m_\alpha m_\mu}. \quad (8.39)$$

Так же как и в предыдущих случаях, интегрирование по $\boldsymbol{k}$ дает нуль, если $\beta\neq\lambda$. Мы можем опять перейти от интегрирования по $\boldsymbol{k}$ к интегрированию по ε и воспользоваться обозначением (8.13), тогда

$$\sigma^{(i)}_{\lambda\mu\nu\rho}=-\frac{e^4}{c^2}n^{(i)}\langle\tau^3\rangle\sum_\alpha\frac{\delta_{\alpha\mu\nu}\delta_{\alpha\lambda\rho}}{m_\alpha m_\mu m_\lambda}. \quad (8.40)$$

Так как замена $\nu\rightleftarrows\rho$ не меняет члена $E_\mu H_\nu H_\rho$, то $\sigma^{(i)}_{\lambda\mu\nu\rho}$ в сумме (8.38) удобно симметризировать, полагая

$$(\sigma^{(i)}_{\lambda\mu\nu\rho})_{\text{симм}}=\tfrac{1}{2}(\sigma^{(i)}_{\lambda\mu\nu\rho}+\sigma^{(i)}_{\lambda\mu\rho\nu}). \quad (8.41)$$

В этом случае каждому коэффициенту $(\sigma^{(i)}_{\lambda\mu\nu\rho})_{\text{симм}}$ соответствует одно определенное произведение $E_\mu H_\nu H_\rho$.

В дальнейшем мы будем предполагать в (8.38) такую симметризацию; тогда, опуская знак $(\quad)_{\text{симм}}$, получим вместо (8.40)

$$\sigma^{(i)}_{\lambda\mu\nu\rho}=\frac{e^4}{c^2}n^{(i)}\langle\tau^3\rangle M^{(i)}_{\lambda\mu\nu\rho}, \quad (8.42)$$

где

$$M^{(i)}_{\lambda\mu\nu\rho}=-\frac{1}{2}\sum_\alpha\frac{1}{m_\alpha m_\lambda m_\mu}[\delta_{\alpha\mu\nu}\delta_{\alpha\lambda\rho}+\delta_{\alpha\mu\rho}\delta_{\alpha\lambda\nu}]. \quad (8.42а)$$

Отсюда следует, что

$$M^{(i)}_{\lambda\lambda\lambda\lambda}=0, \quad (8.43)$$

$$M^{(i)}_{\lambda\lambda\nu\nu}=-\frac{1}{m_\lambda^2 m_\alpha}\quad(\alpha\neq\lambda,\ \nu), \quad (8.43а)$$

$$M^{(i)}_{\lambda\mu\lambda\mu}=M^{(i)}_{\lambda\mu\mu\lambda}=\frac{1}{2m_1m_2m_3}. \quad (8.43б)$$

Коэффициенты $\sigma^{(i)}_{\lambda\mu\nu\rho}$ учитывают изменение тока, связанное с изменением электрического сопротивления (магнетосопротивление). Из (8.42) и (8.43) следует, что $\sigma^{(i)}_{\lambda\lambda\lambda\lambda}=0$, т. е. магнетосопротивление вдоль одной из главных осей эллипсоида энергии, в том случае, когда направление магнитного поля совпадает с направлением тока, равно нулю. Этот результат обобщает выводы об исчезновении продольного магнетосопротивления в случае сферических поверхностей энергии (§ 5, п.1).

Для германия и кремния все разные и отличные от нуля коэффициенты $M^{(i)}_{\lambda\mu\nu\rho}$ равны

$$M^{(i)}_{1122}=-\frac{1}{m_\perp^2 m_\parallel},\quad M^{(i)}_{1133}=M^{(i)}_{2233}=-\frac{1}{m_\perp^3},$$

$$M^{(i)}_{3311}=M^{(i)}_{3322}=-\frac{1}{m_\parallel^2 m_\perp},\quad M^{(i)}_{1212}=M^{(i)}_{1313}=M^{(i)}_{2323}=\frac{1}{2m_\perp^2 m_\parallel}. \quad (8.44)$$

Конечно, имеют место все соотношения вида: $M^{(i)}_{1122}=M^{(i)}_{2211}$ или $M^{(i)}_{1212}=M^{(i)}_{1221}$ и т. д.

В кремнии и германии совокупность всех эквивалентных минимумов в первой бриллюэновской зоне удовлетворяет кубической симметрии кристалла. Мы различаем два случая:

А. Шесть эллипсоидов энергии расположены вдоль направления ⟨100⟩ (кремний).

Б. Четыре эллипсоида энергии расположены вдоль направлений ⟨111⟩ (германий).

Обозначим через x_i ($i=1, 2, 3$) оси прямоугольной координатной системы, направленные вдоль ребер куба элементарной ячейки, и через ξ_i ($i=1, 2, 3$) главные оси одного из эллипсоидов энергии.

В случае А для одной пары эллипсоидов $\xi_1 \parallel x_1$, $\xi_2 \parallel x_2$, $\xi_3 \parallel x_3$, для другой пары — $\xi_1 \parallel x_1$, $\xi_2 \parallel x_3$, $\xi_3 \parallel x_2$ и для третьей пары — $\xi_1 \parallel x_3$, $\xi_2 \parallel x_2$, $\xi_3 \parallel x_1$.

Если токи от отдельных эллипсоидов просто складываются, то, например,

$$\sigma_{1122}=\sum_i \sigma^{(i)}_{\lambda\mu\nu\rho}=2\sigma^{(i)}_{1122}+2\sigma^{(i)}_{1133}+2\sigma^{(i)}_{3322}=n\frac{e^4}{c^2}\langle\tau^3\rangle M_{1122}, \quad (8.45)$$

где концентрация электронов $n=N_c n^{(i)}$, $N_c=6$ — число эквивалентных минимумов и

$$M_{1122}=\frac{1}{N_c}\sum_i M^{(i)}_{\lambda\mu\nu\rho}=\frac{1}{3}\left[M^{(i)}_{1122}+M^{(i)}_{1133}+M^{(i)}_{3322}\right]=$$

$$=-\frac{1}{3}\frac{m_\perp^2+m_\perp m_\parallel+m_\parallel^2}{m_\perp^3 m_\parallel}. \quad (8.45а)$$

Аналогично могут быть вычислены все отличные от нуля коэффициенты $M_{\lambda\mu\nu\rho}$ в обоих случаях А и Б. В случае Б дело не-

сколько осложняется тем, что главные оси эллипсоидов энергии не параллельны ребрам куба элементарной ячейки, поэтому для определения компонент тензора $M_{\lambda\mu\nu\rho}$ через компоненты тензоров $M^{(i)}_{\lambda\mu\nu\rho}$ надо преобразовать последние к координатной системе O, связанной с кристаллом (см. Приложение 11).

Для сравнения теории с опытом представляется более удобным ввести вместо величин $M^{(i)}_{\lambda\mu\nu\rho}$ безразмерные коэффициенты

$$F^{(i)}_{\lambda\mu\nu\rho} = \frac{m''^4}{m'} M^{(i)}_{\lambda\mu\nu\rho}, \tag{8.46}$$

зависящие только от отношения компонент тензора эффективной массы. Используя (8.16б), (8.31б) и (8.44), легко составить

Таблица IX.1

$F^{(i)}_{1122}$	$-\dfrac{3\,(m_\perp + 2m_\parallel)\,m_\perp}{(m_\parallel + 2m_\perp)^2}$
$F^{(i)}_{1133} = F^{(i)}_{2233}$	$-\dfrac{3\,(m_\perp + 2m_\parallel)\,m_\parallel}{(m_\parallel + 2m_\perp)^2}$
$F^{(i)}_{3311} = F^{(i)}_{3322}$	$-\dfrac{3\,(m_\perp + 2m_\parallel)\,m_\perp^2}{(m_\parallel + 2m_\perp)^2\, m_\parallel}$
$F^{(i)}_{1212} = F^{(i)}_{1313} = F^{(i)}_{2323}$	$\dfrac{3\,(m_\perp + 2m_\parallel)\,m_\perp}{2\,(m_\parallel + 2m_\perp)^2}$

табл. IX.1. Компоненты тензора магнетопроводимости $\sigma_{\lambda\mu\nu\rho}$ в координатной системе, связанной с осями кристалла, равны

$$\sigma_{\lambda\mu\nu\rho} = \sum_i \sigma^{(i)}_{\lambda\mu\nu\rho} = \frac{e^4}{c^2}\langle\tau^3\rangle \frac{m'}{m''^4} \sum_i n^{(i)} F^{(i)}_{\lambda\mu\nu\rho} =$$

$$= \frac{e^4}{c^2}\langle\tau^3\rangle \frac{m'}{m''^4}\, n\, \frac{1}{N_c} \sum_i F^{(i)}_{\lambda\mu\nu\rho} = \frac{\langle\tau^3\rangle}{\langle\tau\rangle^3}\left(\sigma\,\frac{\mu^2}{c^2}\right)\left(\frac{m'}{m}\right)^4 F_{\lambda\mu\nu\rho}, \tag{8.47}$$

где суммирование производится по всем эквивалентным минимумам бриллюэновской зоны. Здесь σ и μ — электропроводность и подвижность при $H = 0$ (8.16) и (8.17), N_c — число эквивалентных минимумов в бриллюэновской зоне и

$$F_{\lambda\mu\nu\rho} = \frac{1}{N_c} \sum_i F^{(i)}_{\lambda\mu\nu\rho}. \tag{8.47а}$$

При вычислении $F_{\lambda\mu\nu\rho}$, аналогичном вычислению $M_{\lambda\mu\nu\rho}$, необходимо компоненты тензоров $F^{(i)}_{\lambda\mu\nu\rho}$ относящихся к различным

минимумам, преобразовать от главных осей эллипсоидов энергии к координатной системе, связанной с осями кристалла. В случае А это делается элементарно; в случае Б необходимо составить таблицу направляющих косинусов между главными осями эллипсоидов энергии и осями, совпадающими с ребрами куба элементарной ячейки. Результаты таких вычислений сведены в табл. IX.2. Запишем электрический ток, соответствующий компоненте тензора $\sigma_{\lambda\mu\nu\rho}$ (8.47), в виде

$$\frac{\langle\tau^3\rangle}{\langle\tau\rangle^3}\times\left(\frac{m'}{m''}\right)^4 F_{\lambda\mu\nu\rho}\times\left(\sigma E_\mu\frac{\mu H_\nu}{c}\frac{\mu H_\rho}{c}\right). \tag{8.48}$$

Здесь первый безразмерный множитель порядка 1 связан с зависимостью τ от энергии ε; если τ от ε не зависит, он равен 1. Второй безразмерный множитель связан только с отношением компонент тензора эффективной массы и конфигурацией эллипсоидов энергии в бриллюэновской зоне. Наконец, третий множитель равен омическому току σE_μ, умноженному на два малых безразмерных множителя, пропорциональных магнитному полю.

Из табл. IX.2 видно, что в общем случае симметричный тензор $\sigma_{\lambda\mu\nu\rho}$ в кубическом кристалле описывается тремя независимыми величинами: $\sigma_{\lambda\lambda\lambda\lambda}$, $\sigma_{\lambda\lambda\mu\mu}$, $\sigma_{\lambda\mu\lambda\mu}+\sigma_{\lambda\mu\mu\lambda}$. Соответствующий этому тензору ток, например в направлении оси 1, равен

$$j_1=\sum_{\mu\nu\rho}\sigma_{1\mu\nu\rho}E_\mu H_\nu H_\rho=\sigma_{1111}E_1H_1^2+\sigma_{1122}E_1H_2^2+\sigma_{1133}E_1H_3^2+$$
$$+(\sigma_{1212}+\sigma_{1221})E_2H_1H_2+(\sigma_{1313}+\sigma_{1331})E_3H_1H_3. \tag{8.49}$$

Симметризуя подобно (8.41) коэффициенты и обозначая их так же, получим

$$j_1=\sigma_{1111}E_1H_1^2+\sigma_{1122}E_1H_2^2+\sigma_{1133}E_1H_3^2+2\sigma_{1212}E_2H_1H_2+2\sigma_{1313}E_3H_1H_3. \tag{8.49a}$$

Обозначая

$$\sigma_{1111}=\gamma',\quad \sigma_{1212}=\sigma_{1313}=\alpha,\quad 2\sigma_{1212}=2\sigma_{1313}=\beta, \tag{8.50}$$

запишем (8.49а) в виде

$$j_1=\alpha E_1(H_1^2+H_2^2+H_3^2)+\beta H_1(E_1H_1+E_2H_2+E_3H_3)+\gamma E_1H_1^2, \tag{8.51}$$

где

$$\gamma=\gamma'-\alpha-\beta. \tag{8.51a}$$

Из выражения (8.51) видно, что плотность соответствующего тока

$$\boldsymbol{j}=\alpha H^2\boldsymbol{E}+\beta(\boldsymbol{EH})\boldsymbol{H}+\gamma\boldsymbol{F}, \tag{8.52}$$

где вектор

$$\boldsymbol{F}(E_1H_1^2,\ E_2H_2^2,\ E_3H_3^2), \tag{8.52a}$$

т. е. имеет по оси μ составляющую $E_\mu H_\mu^2$.

Учитывая (8.32), видим, что полный ток

$$\boldsymbol{j}=\sigma\boldsymbol{E}+\eta[\boldsymbol{EH}]+\alpha H^2\boldsymbol{E}+\beta(\boldsymbol{EH})\boldsymbol{H}+\gamma\boldsymbol{F}, \tag{8.53}$$

Таблица IX.2

	$F_{\lambda\lambda\lambda\lambda}$	$F_{\lambda\lambda\nu\nu}$	$F_{\lambda\mu\lambda\mu}=F_{\lambda\mu\mu\lambda}$
⟨100⟩	0	$-\dfrac{\left(1+2\dfrac{m_\parallel}{m_\perp}\right)\left[1+\dfrac{m_\parallel}{m_\perp}+\left(\dfrac{m_\parallel}{m_\perp}\right)^2\right]}{\dfrac{m_\parallel}{m_\perp}\left(2+\dfrac{m_\parallel}{m_\perp}\right)^2}$	$\dfrac{3\left(1+2\dfrac{m_\parallel}{m_\perp}\right)}{2\left(2+\dfrac{m_\parallel}{m_\perp}\right)^2}$
⟨111⟩	$-\dfrac{2\left(\dfrac{m_\parallel}{m_\perp}-1\right)^2\left(1+2\dfrac{m_\parallel}{m_\perp}\right)}{3\dfrac{m_\parallel}{m_\perp}\left(2+\dfrac{m_\parallel}{m_\perp}\right)^2}$	$-\dfrac{\left(1+2\dfrac{m_\parallel}{m_\perp}\right)\left[2+5\dfrac{m_\parallel}{m_\perp}+2\left(\dfrac{m_\parallel}{m_\perp}\right)^2\right]}{3\dfrac{m_\parallel}{m_\perp}\left(2+\dfrac{m_\parallel}{m_\perp}\right)^2}$	$\dfrac{\left(1+2\dfrac{m_\parallel}{m_\perp}\right)\left[2+5\dfrac{m_\parallel}{m_\perp}+2\left(\dfrac{m_\parallel}{m_\perp}\right)^2\right]}{6\dfrac{m_\parallel}{m_\perp}\left(2+\dfrac{m_\parallel}{m_\perp}\right)^2}$
	$\sigma_{\lambda\lambda\lambda\lambda}/\sigma$	$\sigma_{\lambda\lambda\nu\nu}/\sigma$	$\sigma_{\lambda\mu\lambda\mu}/\sigma=\sigma_{\lambda\mu\mu\lambda}/\sigma$
	$-(a+b+c)$	$-a-(R\sigma)^2$	$-\dfrac{1}{2}b+\dfrac{1}{2}(R\sigma)^2$

т. е. характеризуется пятью константами σ, η, α, β и γ. В изотропном случае, т. е. при сферической форме поверхностей энергии,

$$\alpha=\beta \quad \text{и} \quad \gamma=0, \tag{8.53а}$$

как это следует из симметрии.

Общий вид (8.53) с неопределенными константами σ, η, α, β, γ для кубических кристаллов может быть получен на основании одних соображений симметрии. Отсюда следует, что любая правильная электронная теория для тока $\boldsymbol{j}$ в приближении H^2 должна для кристаллов кубической симметрии приводить к выражению вида (8.53).

При экспериментальном изучении магнетосопротивления исследуется не зависимость тока $\boldsymbol{j}$ от полей $\boldsymbol{E}$ и $\boldsymbol{H}$, а разность потенциалов (т. е. электрическое поле $\boldsymbol{E}$) в зависимости от $\boldsymbol{j}$ и $\boldsymbol{H}$; поэтому желательно выражение (8.53) решить относительно $\boldsymbol{E}$. Можно, однако, показать, что из одних соображений симметрии для кубического кристалла, аналогично (8.53),

$$\boldsymbol{E}=\rho\boldsymbol{j}+g[\boldsymbol{jH}]+aH^2\boldsymbol{j}+b(\boldsymbol{jH})\boldsymbol{H}+c\boldsymbol{L}, \tag{8.54}$$

где ρ, g, b, c — постоянные, а вектор $\boldsymbol{L}$ имеет составляющие $j_1H_1^2$, $j_2H_2^2$ и $j_3H_3^2$. Выражения (8.53) и (8.54) имеют аналогичную структуру.

Подставляя (8.54) в (8.53), учитывая члены порядка не выше H^2 и сравнивая коэффициенты при $\boldsymbol{j}$, $[\boldsymbol{jH}]$, $H^2\boldsymbol{j}$, $(\boldsymbol{jH})$ и $\boldsymbol{L}$, получим простую систему уравнений для ρ, g, a, b, c, откуда

$$\left.\begin{aligned} &\rho=\frac{1}{\sigma}, \quad g=-\frac{\eta}{\sigma}, \quad a=-\frac{(\alpha+\eta^2/\sigma)}{\sigma}, \\ &b=-\frac{(\beta-\eta^2/\sigma)}{\sigma}, \quad c=-\frac{\gamma}{\sigma}. \end{aligned}\right\} \tag{8.55}$$

В отсутствие магнитного поля направление электрического поля $\boldsymbol{E}$ совпадает с направлением тока $\boldsymbol{j}$, поэтому удельное сопротивление $\rho=1/\sigma=E/j$. В присутствии магнитного поля, когда $\boldsymbol{E}$ не параллельно $\boldsymbol{j}$, магнетосопротивление $\rho_H=E_j/j=(\boldsymbol{Ej})/j^2$, где E_j — проекция электрического поля на направление тока. Используя (8.54), получим

$$\rho_H=\frac{\boldsymbol{Ej}}{j^2}=\rho+\rho aH^2+\rho b\frac{(\boldsymbol{jH})^2}{j^2}+\rho c\frac{\boldsymbol{Lj}}{j^2}. \tag{8.56}$$

Отсюда

$$\frac{\Delta\rho}{\rho}=\frac{\rho_H-\rho}{\rho}=aH^2+\frac{b(j_1H_1+j_2H_2+j_3H_3)^2}{j^2}+c\frac{j_1^2H_1^2+j_2^2H_2^2+j_3^2H_3^2}{j^2}. \tag{8.56а}$$

Величины a, b, c могут быть определены из опыта, они связаны посредством (8.55) с σ, η (или R), α, β, γ, которые вычисляются из теории, как это показано выше. В табл. IX.2 отношения $\sigma_{\lambda\mu\nu\rho}/\sigma$ выражены через измеримые на опыте константы a, b, c и $R\sigma$.

Из предыдущего видно, что измерения эффекта Холла и магнетосопротивления в монокристаллах полупроводника при различных направлениях тока и магнитного поля дают важную информацию об эффективных массах носителей тока и расположении энергетических поверхностей в бриллюэновской зоне. Такие измерения в n-германии и n-кремнии, приведенные до исследования в них циклотронного резонанса, привели к правильной картине энергетического спектра электронов проводимости.

Как выше отмечалось, из (8.42) и (8.43) следует, что если ток и магнитное поле параллельны главной оси эллипсоида энергии, то продольное магнетосопротивление, аналогично случаю сферических поверхностей энергии, равно нулю. При изучении продольного магнетосопротивления в кремнии было обнаружено, что если ток и параллельное ему магнитное поле совпадают с направлением <100>, то магнетосопротивление падает почти до нуля. В то же время изучение продольного магнетосопротивления в германии показало, что не существует в кристалле такого направления, вдоль которого оно сильно уменьшалось бы. Отсюда было сделано заключение, что в кремнии эквивалентные минимумы расположены вдоль направлений <100>, а в германии — вдоль направлений <111>.

Мы не будем рассматривать здесь для многоэллипсоидной модели случай сильных магнитных полей, разобранный в специальной литературе [1]). Отметим только, что в сильных полях постоянная Холла и в многоэллипсоидной модели равна

$$R = -1/cen,$$

аналогично тому, как это имело место во всех случаях изотропной модели (5.16). Это связано с тем, что в случае сильного магнитного поля холловский ток не связан с рассеянием электронов, а определяется их свободным дрейфом в направлении, перпендикулярном к $\boldsymbol{E}$ и $\boldsymbol{H}$ со скоростью $(E/H)c$.

6. Для определения дифференциальной термоэдс необходимо исходить из уравнения (8.4), в котором надо положить $\boldsymbol{H} = 0$; тогда

$$\boldsymbol{v}\nabla_r f - \frac{e}{\hbar}\boldsymbol{E}\nabla_k f = -\frac{f - f_0}{\tau(\varepsilon)}. \tag{8.57}$$

С точностью до членов, линейных в $\boldsymbol{E} = -\nabla\varphi$ (φ — электростатический потенциал) и ∇T, можно в левую часть (8.57) подставить $f = f_0(\varepsilon)$; тогда

$$\nabla_r f \approx \nabla_r f_0 = \left(-\frac{\partial f_0}{\partial \varepsilon}\right)\left\{\frac{\varepsilon - \zeta}{T}\nabla T + \nabla\zeta\right\}, \tag{8.58}$$

[1]) Abeles B., Meiboom S. — Phys. Rev., 1954, v. 95, p. 31; Shibuya M. — Phys. Rev. 1954, v. 95, p. 1385.

так как при градиенте температуры от координат зависит T и $\zeta(T)$.

Учитывая (8.58) и (8.9), получим из (8.57)

$$\left(-\frac{\partial f_0}{\partial \varepsilon}\right) \boldsymbol{v}\left[\frac{\varepsilon-\zeta}{T}\nabla T - e\nabla\left(\varphi-\frac{\zeta}{e}\right)\right] = -\frac{f-f_0}{\tau}. \qquad (8.59)$$

Электрический ток

$$\boldsymbol{j} = -e\sum_{(\boldsymbol{k})}(f-f_0)\,\boldsymbol{v} =$$

$$= e\sum_{(\boldsymbol{k})}\left(-\frac{\partial f_0}{\partial \varepsilon}\right)\tau(\varepsilon)\left\{\boldsymbol{v}\left[\frac{\varepsilon-\zeta}{T}\nabla T - e\nabla\left(\varphi-\frac{\zeta}{e}\right)\right]\right\}\boldsymbol{v}, \qquad (8.60)$$

где суммирование $\sum_{(\boldsymbol{k})}$ (интегрирование) ведется по волновому вектору $\boldsymbol{k}$. Для тока $\boldsymbol{j}=0$ получим из (8.60)

$$k_0\sum_{(\boldsymbol{k})}\left(-\frac{\partial f_0}{\partial \varepsilon}\right)\tau(\varepsilon)\frac{\varepsilon-\zeta}{k_0T}(\boldsymbol{v}\nabla T)\,\boldsymbol{v} = e\sum_{(\boldsymbol{k})}\left(-\frac{\partial f_0}{\partial \varepsilon}\right)\tau(\varepsilon)\left(\boldsymbol{v}\nabla\left(\varphi-\frac{\zeta}{e}\right)\right)\boldsymbol{v}. \qquad (8.61)$$

Если ∇T направлен вдоль оси μ, то $\boldsymbol{v}\nabla T = v_\mu \nabla_\mu T$ и выражение, стоящее слева в (8.61), равно

$$\nabla_\mu T k_0\sum_{(\boldsymbol{k})}\left(-\frac{\partial f_0}{\partial \varepsilon}\right)\tau\frac{\varepsilon-\zeta}{k_0T}v_\mu^2 = \nabla_\mu T k_0\frac{1}{m_\mu}\frac{2}{3}\sum_{(\varepsilon)}\left(-\frac{\partial f_0}{\partial \varepsilon}\right)\tau\frac{\varepsilon-\zeta}{k_0T}\varepsilon =$$

$$= \nabla_\mu T k_0\frac{1}{m_\mu}\left\{\frac{\langle\tau\varepsilon\rangle}{k_0T} - \frac{\langle\tau\zeta\rangle}{k_0T}\right\}, \qquad (8.62)$$

где мы воспользовались соображениями, изложенными при переходе к (8.12б), (8.12в) и (8.13). Очевидно, что вектор правой части (8.61) должен быть параллелен $\nabla_\mu T$, так что правая часть равна

$$\frac{e}{m_\mu}\nabla_\mu\left(\varphi-\frac{\zeta}{e}\right)\langle\tau\rangle. \qquad (8.63)$$

Из (8.62) и (8.63) следует, что дифференциальная термоэдс

$$\alpha = \frac{\nabla_\mu\left(\varphi-\frac{\zeta}{e}\right)}{\nabla_\mu T} = \left(\frac{k_0}{e}\right)\left\{\frac{\langle\tau\varepsilon\rangle}{k_0T\langle\tau\rangle} - \frac{\zeta}{k_0T}\right\}. \qquad (8.64)$$

Мы видим, что первый член в фигурных скобках совпадает с тем, что дает изотропная модель (сферические поверхности энергии); что касается химического потенциала, то он равен

$$\frac{\zeta}{k_0T} = \ln\frac{4\pi^3\hbar^3 n}{(2\pi k_0T)^{3/2}(m_1m_2m_3)^{1/2}N_c}, \qquad (8.64а)$$

как это следует из (VI.2.12б) и (VI.2.22а).

7. В заключение мы кратко остановимся на упрощенной теории явлений переноса в полупроводниках типа *p*-германий.

В гл. IV, § 15 мы описали структуру валентной зоны p-германия и p-кремния. Наиболее существенным для нее является двукратное вырождение энергетического спектра дырок при $\boldsymbol{k}=0$ и связанное с этим наличие двух сортов дырок: тяжелых и легких. Было отмечено, что при изотропном усреднении по $\boldsymbol{k}$-пространству энергия дырок становится пропорциональной $|\boldsymbol{k}|^2$. Отношение скалярных эффективных масс тяжелых и легких дырок в германии: $m_1/m_2=8{,}0$.

В таком изотропном приближении для энергий и в предположении, что времена релаксации τ_1 и τ_2 для тяжелых и легких дырок зависят от их энергий, нетрудно вычислить все кинетические коэффициенты.

Сравнение теории с опытом показывает, что для германия отношение подвижностей легких и тяжелых дырок примерно равно обратному отношению их эффективных масс [1])

$$\mu_2/\mu_1 \approx m_1/m_2=8{,}0. \tag{8.65}$$

Если τ_1 и τ_2 были бы много меньше, чем времена перехода из состояния тяжелых дырок в состояние легких дырок (мы будем называть эти переходы межзонными), то времена релаксаций можно было бы вычислять обычным способом, т. е. без учета межзонных переходов, и в этом случае, согласно (VIII.4.11), для взаимодействия с акустическими колебаниями $\frac{\langle\tau_1\rangle}{\langle\tau_2\rangle}=\left(\frac{m_2}{m_1}\right)^{3/2}$. При этом мы предполагаем, что константа деформационного потенциала $\mathscr{E}_1$ (VIII.5.15) (или константа взаимодействия C) одна и та же для тяжелых и легких дырок. Последнее может быть обосновано тем, что $\mathscr{E}_1$, согласно (VIII.5.4), равно смещению края валентной зоны, отнесенному к относительному изменению объема при равномерном сжатии кристалла. Так как при равномерном сжатии вырождение не снимается, то значение $\mathscr{E}_1$ одинаково для обоих сортов дырок.

Используя обычную формулу для подвижности (3.3), получим $\frac{\mu_2}{\mu_1}=\left(\frac{m_1}{m_2}\right)^{5/2}$, что находится в противоречии с (8.65). Это показывает, что при рассеянии должны играть существенную роль межзонные переходы. Из (8.65) следует, что времена релаксаций тяжелых и легких дырок должны быть одинаковы.

Определим время релаксации, например для тяжелых дырок, с учетом межзонных переходов при упрощающем предположении, что вероятность перехода при рассеянии на акустических колебаниях и в случае вырожденной валентной зоны имеет простую изотропную форму (VIII.3.11а), (VIII.3.11б) [2]).

---

[1]) Willardson R. K., Harman T. C., Beer A. C.—Phys. Rev., 1954, v. 96, p. 1512.

[2]) Пикус Г. Е.—ЖТФ, 1957, т. 27, с. 1606.

Считая столкновения с фононами упругими, заменяя N_q и N_q+1 на $k_0T/\hbar\omega_q$ и полагая для акустической ветви $\omega_q=v_0q$ (v_0—скорость продольных волн), получим как для испускания, так и для поглощения фонона одинаковую вероятность

$$W(\boldsymbol{k}_1\boldsymbol{k}')=W_0\delta(\varepsilon_{k'}-\varepsilon_k), \tag{8.66}$$

где

$$W_0=\frac{4\pi C^2k_0T}{9NM\hbar v_0^2}. \tag{8.66а}$$

Обозначим функции распределения тяжелых и легких дырок через $f_1(\boldsymbol{k})$ и $f_2(\boldsymbol{k})$, тогда изменение f_1 из-за столкновений, согласно (VIII.2.8), равно

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{ст}}=-W_0\sum_{k'}\{[f_1(\boldsymbol{k})-f_1(\boldsymbol{k}')]\,\delta[\varepsilon_1(\boldsymbol{k})-\varepsilon_1(\boldsymbol{k}')]+ \\ +[f_1(\boldsymbol{k})-f_2(\boldsymbol{k}')]\,\delta[\varepsilon_1(\boldsymbol{k})-\varepsilon_2(\boldsymbol{k}')]\}, \tag{8.67}$$

где

$$\varepsilon_1(\boldsymbol{k})=\hbar^2k^2/2m_1,\quad \varepsilon_2(\boldsymbol{k})=\hbar^2k^2/2m_2 \tag{8.67а}$$

— энергии тяжелых и легких дырок.

Первая квадратная скобка в (8.67) учитывает переходы тяжелых дырок в тяжелые (из состояния $\boldsymbol{k}$ в $\boldsymbol{k}'$), во втором слагаемом — переходы тяжелых дырок в легкие.

Положим, как обычно,

$$f_1(\boldsymbol{k})=f_{10}(\varepsilon)+\boldsymbol{g}_1(\varepsilon)\boldsymbol{k}, \\ f_2(\boldsymbol{k})=f_{20}(\varepsilon)+\boldsymbol{g}_2(\varepsilon)\boldsymbol{k}. \tag{8.68}$$

Подставляя $f_1(\boldsymbol{k})$ в первую квадратную скобку (8.67) и производя интегрирование по $\boldsymbol{k}'$, получим результат, совпадающий с (гл. VIII, § 4)

$$\left(\frac{\partial f_1}{\partial t}\right)^{(11)}_{\text{ст}}=-\frac{W_0(2\varepsilon)^{1/2}}{\pi^2\hbar^3}m_1^{3/2}(\boldsymbol{g}_1\boldsymbol{k})=-\frac{(\boldsymbol{g}_1\boldsymbol{k})}{\tau_{11}}, \tag{8.69}$$

где τ_{11}—время релаксации для тяжелых дырок, обусловленное их переходами в той же зоне.

Заметим, что

$$f_{10}(\varepsilon)=f_{20}(\varepsilon)=e^{\frac{\zeta-\varepsilon}{k_0T}}, \tag{8.70}$$

так как в равновесии химический потенциал для тяжелых и легких дырок одинаков.

Подставляя $f_1(\boldsymbol{k})$ и $f_2(\boldsymbol{k})$ во вторую квадратную скобку (8.67) и используя (8.70), легко убедиться в том, что при интегрировании по $\boldsymbol{k}'$ член, пропорциональный $\boldsymbol{g}_2(\varepsilon)\boldsymbol{k}'$, равен нулю, поэтому

$$\left(\frac{\partial f_1}{\partial t}\right)^{(12)}_{\text{ст}}=-\frac{W_0(2\varepsilon)^{1/2}}{\pi^2\hbar^3}m_2^{3/2}(\boldsymbol{g}_1\boldsymbol{k})=-\frac{(\boldsymbol{g}_1\boldsymbol{k})}{\tau_{12}}, \tag{8.71}$$

где τ_{12} — время релаксаций для перехода тяжелых дырок в легкие. В результате полное время релаксаций для тяжелых дырок

$$\frac{1}{\tau_1}=\frac{1}{\tau_{11}}+\frac{1}{\tau_{12}}=\frac{W_0\,(2\varepsilon)^{1/2}}{\pi^2\hbar^3}\,(m_1^{3/2}+m_2^{3/2}). \tag{8.72}$$

Легко видеть, что аналогичный расчет для легких дырок дает для обратного времени релаксации $1/\tau_2$ выражение, совпадающее с (8.72).

Таким образом, в принятом нами приближении изотропной вероятности рассеяния (8.66) и сферической формы поверхностей энергии (8.67а) времена релаксаций тяжелых и легких дырок одинаковы, т. е.

$$\tau_1=\tau_2, \tag{8.73}$$

откуда непосредственно следует (8.65).

Электропроводность, обусловленная обоими сортами дырок, согласно (3.3), равна

$$\sigma=\frac{n_1e^2}{m_1}\langle\tau_1\rangle+\frac{n_2e^2}{m_2}\langle\tau_2\rangle=n_1e\mu_1+n_2e\mu_2. \tag{8.74}$$

Здесь n_i и μ_i — концентрации и подвижности тяжелых и легких дырок.

Для определения постоянной Холла и магнетосопротивления воспользуемся уравнениями (5.4), где a_1 и a_2 определяются токами как тяжелых, так и легких дырок:

$$\begin{aligned}a_1&=\frac{n_1e^2}{m_1}\left\langle\frac{\tau_1}{1+(\gamma_1\tau_1H)^2}\right\rangle+\frac{n_2e^2}{m^2}\left\langle\frac{\tau_2}{1+(\gamma_2\tau_2H)^2}\right\rangle,\\ a_2&=-\frac{n_1e^2}{m_1}(\gamma_1H)\left\langle\frac{\tau_1^2}{1+(\gamma_1\tau_1H)^2}\right\rangle-\frac{n_2e^2}{m_2}(\gamma_2H)\left\langle\frac{\tau_2^2}{1+(\gamma_2\tau_2H)^2}\right\rangle.\end{aligned} \tag{8.75}$$

Здесь оба слагаемых в a_2 отрицательны, так как у нас по-прежнему $\gamma_1=\frac{e}{m_{1c}}>0$ и $\gamma_2=\frac{e}{m_{2c}}>0$, а для электронов a_2 положительно.

Если мы для a_1 и a_2 (8.75) возьмем линейное приближение по H, то по формуле (5.5а) получим для постоянной Холла

$$R=\frac{1}{cen_1}\,\frac{\langle\tau_1^2\rangle+\left(\frac{n_2}{n_1}\right)\left(\frac{m_1}{m_2}\right)^2\langle\tau_2^2\rangle}{\left[\langle\tau_1\rangle+\left(\frac{n_2}{n_1}\right)\left(\frac{m_1}{m_2}\right)\langle\tau_2\rangle\right]^2}. \tag{8.76}$$

Предполагая, что зависимость τ_1 и τ_2 от энергии определяется выражением (3.5а) [1]), легко показать, что

$$\langle\tau^2\rangle = \frac{3\pi}{8}\left(\frac{m}{e}\right)\mu^2, \tag{8.77}$$

где μ — подвижность.

Из (8.76), (8.77) и выражения для подвижности (3.3), получим

$$R = \frac{3\pi}{8}\frac{1}{cen_1}\frac{1+(n_2/n_1)(\mu_2/\mu_1)^2}{[1+(n_2/n_1)(\mu_2/\mu_1)]^2}. \tag{8.76а}$$

Используя для a_1 и a_2 квадратичное приближение вида (5.9), можно вычислить магнетосопротивление в слабых полях. Считая, что дырки имеют время релаксации (3.5а), получим в результате аналогичного расчета

$$\frac{\Delta\rho}{\rho} = \frac{9\pi}{16}\left(\frac{\mu_1 H}{c}\right)^2\left[\frac{1+(n_2/n_1)(\mu_2/\mu_1)^3}{1+(n_2/n_1)(\mu_2/\mu_1)} - \frac{\pi}{4}\left(\frac{1+(n_2/n_1)(\mu_2/\mu_1)^2}{1+(n_2/n_1)(\mu_2/\mu_1)}\right)^2\right]. \tag{8.78}$$

Полагая концентрацию легких дырок $n_2 = 0$, получим из (8.76а) и (8.78) выражения (5.12а) [2]).

Отношение подвижностей легких и тяжелых дырок можно с хорошим приближением считать равным обратному отношению их эффективных масс (8.65). Мы видим, что три выражения (8.74), (8.76а) и (8.78) содержат три неизвестных n_1, μ_1 и n_2/n_1.

Поэтому измерения электропроводности, эффекта Холла и магнетосопротивления, вообще говоря, достаточно для определения концентраций и подвижностей тяжелых и легких дырок.

На самом деле, для сравнения теории с опытом [3]) используются экспериментальные данные и теоретические формулы не только для слабых, но и для промежуточных и сильных магнитных полей.

При этом оказывается, что наилучшее совпадение теории с опытом имеет место, когда отношение концентраций $n_2/n_1 = 0{,}02$, что находится в противоречии с данными об эффективных массах, полученных из циклотронного резонанса, так как, согласно (VI.2.12)

$$\frac{n_2}{n_1} = \left(\frac{m_2}{m_1}\right)^{3/2} = (8{,}0)^{-3/2} = 0{,}04.$$

---

[1]) Хотя германий — типичный атомный кристалл, это предположение нельзя считать обоснованным, так как температурная зависимость подвижности дырок не подчиняется закону (6.9).

[2]) Конечно, тот же результат можно получить, полагая концентрацию тяжелых дырок $n_1 = 0$, однако для этого надо преобразовать (8.76а) и (8.78).

[3]) Willardson R. K., Harman T. C., Beer A. C. — Phys. Rev., 1954, v. 96. p. 1512

Это противоречие, вероятно, связано с рядом упрощающих предположений теории и с сомнительным предположением о зависимости времени релаксации от энергии (см. примечание на предыдущей странице).

§ 9. Явления переноса в полупроводниках со сферической непараболической зоной

1. Рассмотрим явления переноса в полупроводниках с одним сортом носителей тока, энергия которых ε зависит от абсолютной величины волнового вектора $|\boldsymbol{k}|=k$, но в остальном произвольна, т. е. откажемся от предположения, что энергия ε — однородная квадратичная функция составляющих волнового вектора k_x, k_y, k_z. Такая ситуация может возникнуть в полупроводнике с узкой запрещенной зоной и реализуется, например, в антимониде индия (InSb).

Выражение для тока электронов (2.15)[1])

$$\boldsymbol{j}=-\frac{e}{3\pi^2}\int\limits_0^\infty \chi(k)\left(-\frac{\partial f_0}{\partial\varepsilon}\right)u(k)\,k^4dk \tag{9.1}$$

справедливо и в случае нестандартной сферической зоны. Здесь, как следует из (2.2а),

$$u(k)=\frac{1}{\hbar k}\frac{\partial\varepsilon(k)}{\partial k}. \tag{9.1а}$$

Рассмотрим вначале электропроводность в изотермических условиях ($\nabla T=0$) и в отсутствие магнитного поля ($H=0$). В этом случае, из (2.11) следует, что

$$\chi(k)=-eu(k)\,\tau(k)\,\boldsymbol{E}=-\frac{e}{\hbar k}\frac{\partial\varepsilon(k)}{\partial k}\tau(k)\,\boldsymbol{E}. \tag{9.1б}$$

Подставляя (9.1а), (9.1б) в (9.1) получим для электропроводности

$$\sigma=\frac{j}{E}=\frac{e^2}{3\pi^2\hbar^2}\int\limits_0^\infty\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\left(\frac{\partial\varepsilon}{\partial k}\right)^2\tau(k)\,k^4dk. \tag{9.2}$$

Положим

$$\frac{\partial\varepsilon(k)}{dk}=\frac{\hbar^2k}{m(\varepsilon)}, \tag{9.2а}$$

где $m(\varepsilon)$ — эффективная масса, зависящая от энергии электрона ε. В случае стандартной зоны (2.13): $m(\varepsilon)=m$.

Подставим (9.2а) в (9.2) и перейдем от переменной интегрирования k к ε:

$$\sigma=\frac{e^2}{3\pi^2}\int\limits_0^\infty\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\tau(\varepsilon)\frac{k^3(\varepsilon)}{m(\varepsilon)}d\varepsilon. \tag{9.3}$$

[1]) Мы опускаем индекс n.

Для электронов проводимости в InSb энергия, согласно (IV.13.36б), равна

$$\varepsilon = \frac{\hbar^2 k^2}{2m_0} + \frac{1}{2}\varepsilon_G + \frac{(\varepsilon_G^2 + 8P^2k^2/3)^{1/2}}{2}.$$

Здесь m_0 — масса электрона в вакууме, P — константа, характеризующая «взаимодействие» валентной зоны с зоной проводимости, равная (IV.13.30б), ε_G — ширина запрещенной зоны. В написанном выше выражении энергия ε отсчитывается от верхнего края валентной зоны. Для дальнейшего удобнее отсчитывать энергию от нижнего края зоны проводимости; для этого мы должны в написанном выше выражении заменить ε на $\varepsilon + \varepsilon_G$, тогда

$$\varepsilon = \frac{\hbar^2 k^2}{2m_0} + \frac{\varepsilon_G}{2}\left(\sqrt{1 - \frac{8P^2k^2}{3\varepsilon_G^2}} - 1\right). \tag{9.4}$$

Если $Pk \ll \varepsilon_G$, т. е, для малых k, можно разложить корень в ряд, тогда

$$\varepsilon = \frac{\hbar^2 k^2}{2m_0} + \frac{2P^2k^2}{3\varepsilon_G} = \frac{\hbar^2 k^2}{2m(0)}. \tag{9.5}$$

Здесь $m(0)$ — эффективная масса электрона на дне зоны проводимости. Из опытов по циклотронному резонансу известно, что $m(0) = 0{,}013m_0$, т. е. $m(0) \ll m_0$; в этом приближении из (9.4) получим

$$\varepsilon = \frac{\varepsilon_G}{2}\left(\sqrt{1 + \frac{2\hbar^2k^2}{m(0)\varepsilon_G}} - 1\right). \tag{9.6}$$

Отсюда следует, что

$$k(\varepsilon) = \frac{(2m(0)\varepsilon)^{1/2}}{\hbar}\left(1 + \frac{\varepsilon}{\varepsilon_G}\right)^{1/2}. \tag{9.6а}$$

Из (9.2а) и (9.6) получим

$$m(\varepsilon) = m(0)\left(1 + \frac{2\varepsilon}{\varepsilon_G}\right). \tag{9.7}$$

Множитель $(1 + 2\varepsilon/\varepsilon_G)$ в правой части описывает увеличение $m(\varepsilon)$ с ростом ε; при $\varepsilon_G \to \infty$ $m(\varepsilon) = m(0)$ и зона приобретает параболический характер.

Исследуем, как видоизменяется время релаксации $\tau(\varepsilon)$ в случае нестандартной зоны.

Рассмотрим время релаксации, обусловленное взаимодействием электрона с акустическими колебаниями. Оценки показывают, что и в случае нестандартной сферической зоны взаимодействие электронов с акустическими колебаниями происходит почти упруго. Это значит, что в законе сохранения энергии можно пренебречь энергией фонона $\hbar\omega_q = \hbar v_0 q$ (v_0 — скорость звука); легко видеть, что в этом случае (VIII.4.1г): $0 \leqslant q \leqslant 2k$.

Выражение (VIII.4.7) приобретает вид

$$\frac{1}{\tau}=-\frac{V}{(2\pi)^3}\int_0^{2k}q^2\,dq\int_0^{\pi}\sin\vartheta\,d\vartheta\int_0^{2\pi}d\varphi\left\{w(q)\,N_q\delta[\varepsilon((\boldsymbol{k}+\boldsymbol{q})^2)-\right.$$

$$\left.-\varepsilon(\boldsymbol{k}^2)]\frac{q\cos\alpha}{k\cos\beta}-w(q)\,(N_q+1)\,\delta[\varepsilon(\boldsymbol{k}-\boldsymbol{q})^2-\varepsilon(\boldsymbol{k}^2)]\frac{q\cos\alpha}{k\cos\beta}\right\}. \quad (9.8)$$

Здесь $q\cos\alpha$ и $k\cos\beta$ проекции $\boldsymbol{q}$ и $\boldsymbol{k}$ на вектор $\boldsymbol{\chi}$. Очевидно, что δ-функции равны

$$\delta[\varepsilon((\boldsymbol{k}\pm\boldsymbol{q})^2)-\varepsilon(\boldsymbol{k}^2)]=\delta[\varepsilon(k^2+q^2\pm 2kq\cos\vartheta)-\varepsilon(\boldsymbol{k}^2)]=$$

$$=\frac{1}{\left[\pm\frac{\partial\varepsilon}{\partial(k^2)}2kq\right]}\delta\left(\cos\vartheta\mp\frac{q}{2k}\right), \quad (9.9)$$

где $\mp q/2k$ — корни уравнения

$$\varepsilon(k^2+q^2\pm 2kq\cos\vartheta)-\varepsilon(k^2)=0$$

или уравнения

$$k^2+q^2\pm 2kq\cos\vartheta-k^2=0.$$

Действуя дальше так же, как в гл. VIII, § 4, получим

$$\frac{1}{\tau}=\frac{V}{8\pi^2}\frac{1}{k^2}\left(\frac{\partial\varepsilon}{\partial k}\right)^{-1}\int_0^{2k}w(q)\,(2N_q+1)q^3dq, \quad (9.10)$$

откуда

$$\tau=\frac{9\pi}{4}\frac{Mv_0^2\hbar}{\Omega_0c^2k_0T}\left(\frac{\partial\varepsilon}{\partial k}\right)\frac{1}{k^2}. \quad (9.11)$$

В случае стандартной зоны это немедленно приводит к выражению (VIII.4.11). Подставляя в (9.11) (9.2а), (9.7) и (9.6а), получим

$$\tau=\frac{9\pi}{2}\frac{\rho v_0^2\hbar^4}{C^2\,[2m\,(0)\,k_0T]^{3/2}}\left(\frac{\varepsilon}{k_0T}\right)^{-1/2}\frac{\left(1+\frac{\varepsilon}{\varepsilon_G}\right)^{-1/2}}{1+\frac{2\varepsilon}{\varepsilon_G}}. \quad (9.12)$$

Здесь плотность кристалла $\rho=M/\Omega_0$; второй множитель, зависящий от ε_G, учитывает отступление сферической зоны от параболичности. При $\varepsilon_G\to\infty$ выражение (9.12) переходит в (VIII.4.11).

Запишем время релаксации (9.12) в форме

$$\tau=\tau_0x^r\frac{(1+\beta x)^r}{1+2\beta x}, \quad (9.13)$$

где $x=\varepsilon/k_0T$, $\beta=k_0T/\varepsilon_G$ — параметр, характеризующий непараболичность зоны, и $r=-1/2$. Нетрудно показать, что это выражение для времени релаксации применимо и в случае других механизмов рассеяния с другими значениями τ_0 и r (3.5б)—(3.5г). Используя (9.13), (9.6а) и (9.7), получим для

электропроводности (9.3)

$$\sigma=\frac{e^2}{3\pi^2}\frac{[2m(0)\,k_0T]^{3/2}}{\hbar^3 m(0)}\tau_0\int\limits_0^\infty\left(-\frac{\partial f_0}{\partial x}\right)\frac{(x+\beta x^2)^{r+3/2}}{(1+2\beta x)^2}\,dx, \qquad (9.14)$$

где $f_0=[1+\exp(x-z)]^{-1}$, $z=\zeta/k_0T$.

Аналогично могут быть вычислены коэффициенты других кинетических эффектов: термоэлектрических, гальваномагнитных, термомагнитных [1]). Эти коэффициенты выражаются через интегралы, подобные интегралу в (9.14), и имеют вид

$$\mathcal{J}^m_{n,\,k}(z,\ \beta)=\int\limits_0^\infty\left(-\frac{\partial f_0}{\partial x}\right)\frac{x^m(x+\beta x^2)^n}{(1+2\beta x)^k}\,dx; \qquad (9.15)$$

их иногда называют обобщенными или двухпараметрическими интегралами Ферми. При $\beta=0$ они выражаются через интеграл Ферми (VI.2.6); в самом деле,

$$\mathcal{J}^m_{n,\,k}(z,\ 0)=-\int\limits_0^\infty x^{m+n}\frac{\partial f_0}{\partial x}\,dx=-\int\limits_0^\infty x^{m+n}df_0=$$

$$=-x^{m+n}f_0\Big|_0^\infty+(m+n)\int\limits_0^\infty f_0(x)\,x^{m+n-1}dx=(m+n)\,\mathcal{F}_{m+n-1}(z). \qquad (9.15а)$$

Значения табулированных обобщенных интегралов Ферми (9.15) для параметров: $-5\leqslant z\leqslant 20$ и $0\leqslant\beta\leqslant 1$ даны в Приложении Д к книге Б. М. Аскерова [2]).

Выразим концентрацию электронов n через химический потенциал $z=\zeta/k_0T$ и параметр непараболичности $\beta=k_0T/\varepsilon_G$

$$n=2\sum_{\boldsymbol{k}}f_0(\varepsilon)=2\int f_0(\varepsilon)\frac{d^3k}{(2\pi)^3}=\frac{1}{\pi^2}\int f_0(\varepsilon)\,k^2dk=$$

$$=\frac{1}{3\pi^2}\int f_0(\varepsilon)\,dk^3=\frac{1}{3\pi^2}\int\left(-\frac{\partial f_0}{\partial\varepsilon}\right)k^3(\varepsilon)\,d\varepsilon. \qquad (9.16)$$

Подставляя сюда $k(\varepsilon)$ из (9.6а), получим

$$n=\frac{[2m(0)\,k_0T]^{3/2}}{3\pi^2\hbar^3}\int\left(-\frac{\partial f_0}{\partial x}\right)(x+\beta x^2)^{3/2}dx=$$

$$=\frac{[2m(0)\,k_0T]^{3/2}}{3\pi^2\hbar^3}\,\mathcal{J}^0_{3/2,0}(z,\ \beta). \qquad (9.16а)$$

Отсюда и из (9.15а) видно, что для параболической зоны ($\beta=0$)

$$n=\frac{(2mk_0T)^{3/2}}{2\pi^2\hbar^3}\,\mathcal{F}_{1/2}(z), \qquad (9.16б)$$

что совпадает с (VI.2.5).

[1]) Аскеров Б. М. Кинетические эффекты в полупроводниках.—М.: Наука, 1970, гл. V.

[2]) Аскеров Б. М., там же.

В случае сильно вырожденного полупроводника $-\partial f_0/\partial\varepsilon = \delta(\varepsilon-\zeta)$, поэтому из (9.16) следует

$$k(\zeta) = (3\pi^2 n)^{1/3}. \tag{9.17}$$

Отсюда и из (9.6) следует

$$\zeta = \frac{\varepsilon_G}{2}\left[\sqrt{1+\frac{2\hbar^2(3\pi^2 n)^{2/3}}{m(0)\,\varepsilon_G}} - 1\right]. \tag{9.18}$$

Подставляя это в (9.7), получим

$$m(\zeta) = m(0)\sqrt{1+\frac{2\hbar^2(3\pi^2 n)^{2/3}}{m(0)\,\varepsilon_G}}. \tag{9.19}$$

Отсюда следует, что для антимонида индия $m(\zeta)$ возрастает почти в четыре раза, когда концентрация n изменяется от 10^{16} до $5\cdot 10^{18}\,см^{-3}$.

В работе Колодзайчика[1]) приведены для n-InSb зависимость подвижности σ/en от концентрации, которая для параболической зоны вообще отсутствует.

Из (9.18) и (9.19) получаем простую связь между уровнем Ферми ζ и соответствующей ему эффективной массой

$$\zeta = \frac{\varepsilon_G}{2}\left[\frac{m(\zeta)}{m(0)} - 1\right]. \tag{9.20}$$

§ 10. Эффект «фононного увлечения» в полупроводниках

1. Во всех предыдущих параграфах этой главы предполагалось, что фононы находятся в состоянии статистического равновесия, т. е. числа заполнения фононов N_q определяются функцией Планка (VIII.4.10). В то же время очевидно, что при наличии электрического тока и рассеяния электронов на фононах, направленный импульс электронов должен передаваться фононам, в силу чего их распределение не может оставаться изотропным. Это обстоятельство не учитывалось в предыдущих формулах, однако ясно, что явление обратного воздействия неравновесности распределения фононов на неравновесность распределения электронов — эффект 2-го порядка малости и им можно практически пренебречь. Однако отступление распределения фононной функции от равновесного возможно за счет наличия градиента температуры в кристалле.

Как было впервые показано Л. Э. Гуревичем (1945), неравновесность фононного распределения, связанная с наличием градиента температуры, может при определенных условиях играть существенную роль в термоэлектрических явлениях в металлах.

---

[1]) Kolodziejczak J.—Acta Phys. Polon., 1961, v. 20, p. 289.

Это явление получило название эффекта *фононного увлечения* (phonon drag).

Наиболее интересные явления наблюдаются при фононном увлечении в полупроводниках. В 1953 г. Фредериксе и независимо от него Джебл наблюдали значительное увеличение термоэдс α в p-германии при понижении температуры, не предсказываемое обычной теорией.

Херринг и Фредериксе правильно интерпретировали это явление как эффект фононного увлечения дырок в германии. В самом деле, градиент температуры вызывает поток фононов, движущихся от горячего конца образца к холодному, который увлекает за собой носители тока, вызывая тем самым увеличение термоэлектрического тока. Как мы увидим ниже, для правильной оценки величины эффекта увлечения надо учесть, что электроны при рассеянии на акустических колебаниях из-за выполнения законов сохранения могут взаимодействовать только с фононами, волновой вектор которых q порядка волнового вектора электрона k. Поэтому время релаксации фононов, входящее в формулы для фононного увлечения, не совпадает со временем релаксации фононов, определяемым из теплопроводности решетки. Мы рассмотрим теорию фононного увлечения в полупроводниках с одним сортом носителей тока в отсутствие вырождения и в предположении простой зонной структуры (2.13).

2. Количественная теория фононного увлечения основана на кинетическом уравнении [1]). Как мы увидим ниже, эффект фононного увлечения связан, в некотором смысле, с неупругостью рассеяния электронов на акустических колебаниях, которым мы обычно пренебрегаем. Таким образом, при определении $(\partial f/\partial t)_{\text{ст}}$ мы должны исходить из общего выражения (VIII.2.7)

$$\left[\frac{\partial f(\boldsymbol{k})}{\partial t}\right]_{\text{ст}} = \sum_{\boldsymbol{k}'} \{W(\boldsymbol{k}', \boldsymbol{k}) f(\boldsymbol{k}') - W(\boldsymbol{k}, \boldsymbol{k}') f(\boldsymbol{k})\}. \qquad (10.1)$$

Мы не учитываем принцип Паули, так как рассматриваем электроны в невырожденном состоянии. $W(\boldsymbol{k}, \boldsymbol{k}')$ — вероятность перехода из состояния $\boldsymbol{k}$ в $\boldsymbol{k}'$.

Из закона сохранения волнового вектора при взаимодействии с фононом следует

$$\boldsymbol{k}' = \boldsymbol{k} \pm \boldsymbol{q}, \qquad (10.2)$$

если не учитывать процессов переброса. Здесь знак плюс соответствует поглощению, знак минус — испусканию фонона. В соответствии с этим под знаком суммы в (10.1) получим четыре

[1]) Ниже мы следуем способу изложения, предложенному Ю. Н. Образцовым.

слагаемых

$$\left[\frac{\partial f(\boldsymbol{k})}{\partial t}\right]_{\text{ст}} = \sum_{\boldsymbol{q}} \{W^-(\boldsymbol{k}+\boldsymbol{q}, \boldsymbol{k}) f(\boldsymbol{k}+\boldsymbol{q}) + W^+(\boldsymbol{k}-\boldsymbol{q}, \boldsymbol{k}) f(\boldsymbol{k}-\boldsymbol{q}) - W^+(\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{q}) f(\boldsymbol{k}) - W^-(\boldsymbol{k}, \boldsymbol{k}-\boldsymbol{q}) f(\boldsymbol{k})\}, \quad (10.3)$$

где W^+ и W^- — вероятности поглощения и испускания акустических фононов; так что, например, $W^-(\boldsymbol{k}+\boldsymbol{q}, \boldsymbol{k})$ — вероятность перехода $\boldsymbol{k}+\boldsymbol{q} \to \boldsymbol{k}$ в результате испускания фонона с волновым вектором $\boldsymbol{q}$.

Используя (VIII.3.11а) — (VIII.3.11в), запишем (10.3) в виде

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = \sum_{\boldsymbol{q}} w(q) \{f(\boldsymbol{k}+\boldsymbol{q})(N_q+1)\delta(\varepsilon_{\boldsymbol{k}} - \varepsilon_{\boldsymbol{k}+\boldsymbol{q}} + \hbar\omega_q) + f(\boldsymbol{k}-\boldsymbol{q}) N_q \delta(\varepsilon_{\boldsymbol{k}} - \varepsilon_{\boldsymbol{k}-\boldsymbol{q}} - \hbar\omega_q) - f(\boldsymbol{k}) N_q \delta(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{k}} - \hbar\omega_q) - f(\boldsymbol{k})(N_q+1)\delta(\varepsilon_{\boldsymbol{k}-\boldsymbol{q}} - \varepsilon_{\boldsymbol{k}} + \hbar\omega_q)\}. \quad (10.4)$$

Меняя при суммировании во втором и четвертом членах $\boldsymbol{q}$ на $-\boldsymbol{q}$ и используя четность δ-функции, получим

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = \sum_{\boldsymbol{q}} w(q) \{[f(\boldsymbol{k}+\boldsymbol{q})(N_q+1) - f(\boldsymbol{k}) N_q]\delta(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{k}} - \hbar\omega_q) + [f(\boldsymbol{k}+\boldsymbol{q}) N_{-q} - f(\boldsymbol{k})(N_{-q}+1)]\delta(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{k}} + \hbar\omega_q)\}. \quad (10.5)$$

До настоящего времени мы всегда предполагали, что распределение фононов равновесно и в этом случае $N_{-\boldsymbol{q}} = N_{\boldsymbol{q}}$.

Положим в неравновесном случае

$$N_{\boldsymbol{q}} = N_{\boldsymbol{q}}^{(0)} + N'_{\boldsymbol{q}}, \quad (10.6)$$

где $N_{\boldsymbol{q}}^{(0)}$ — равновесная функция распределения фононов.

Как мы сейчас покажем, неравновесная добавка к фононной функции распределения удовлетворяет условию

$$N'_{-\boldsymbol{q}} = -N'_{\boldsymbol{q}}. \quad (10.6а)$$

Неравновесная добавка $N'_{\boldsymbol{q}}$ определяется из уравнений

$$\left(\frac{\partial N'_{\boldsymbol{q}}}{\partial t}\right)_{\text{ст}} = -\frac{N'_{\boldsymbol{q}}}{\tau_{\text{ф}}} \quad (10.7)$$

и

$$\left(\frac{\partial N'_{\boldsymbol{q}}}{\partial t}\right)_{\text{пол}} + \left(\frac{\partial N'_{\boldsymbol{q}}}{\partial t}\right)_{\text{ст}} = 0, \quad (10.7а)$$

где $\tau_{\text{ф}}(q)$ — время релаксации для длинноволновых продольных акустических фононов, а полевой член

$$\left(\frac{\partial N'_{\boldsymbol{q}}}{\partial t}\right)_{\text{пол}} = \boldsymbol{v}_g \nabla_r N_{\boldsymbol{q}}^{(0)}. \quad (10.7б)$$

Здесь $\boldsymbol{v}_g$ — групповая скорость длинноволновых фононов, равная $v_0\frac{\boldsymbol{q}}{q}$, где v_0 — скорость продольных звуковых волн, а

$$\nabla_r N_q^{(0)} = \frac{dN_q^{(0)}}{dT}\nabla T = \frac{k_0}{\hbar v_0 q}\nabla T, \tag{10.7в}$$

так как для длинноволновых фононов $N_q^{(0)} \approx \frac{k_0 T}{\hbar v_0 q}$.

Из (10.7а)—(10.7в) получим

$$N_q' = \tau_\Phi \left(\frac{\partial N_q'}{\partial t}\right)_{\text{пол}} = \frac{k_0 \tau_\Phi}{\hbar q}\left(\frac{\boldsymbol{q}}{q}\nabla T\right), \tag{10.8}$$

откуда следует (10.6а).

Исследования Херринга по теплопроводности кристаллов, на которых мы останавливаться не можем, показали, что если учесть упругую анизотропию кристаллов, устраняющую расходимость в теплопроводности при $q \to 0$, то для продольных длинноволновых фононов в кристаллах кубической симметрии время релаксации для взаимодействия с другими фононами

$$\tau_\Phi(q) = 1/A_l T^3 q^2. \tag{10.9}$$

При этом учет упругой анизотропии кристалла позволяет ограничиться трехфононными столкновениями.

Если мы положим в (10.5) $f(\boldsymbol{k})$ и N_q равными их равновесным значениям, то $\sum\limits_q \{\ \} = 0$. Раньше всегда предполагалось, что $N_q = N_q^{(0)}$, но

$$f(\boldsymbol{k}) = f_0(\varepsilon_{\boldsymbol{k}}) + f_1(\boldsymbol{k}), \tag{10.10}$$

где $f_1(\boldsymbol{k}) \neq 0$, и определялось $(\partial f/\partial t)_{\text{ст}}$, которое при существовании времени релаксации оказывалось пропорциональным $f_1(\boldsymbol{k})$. Подставляя в (10.5) выражения (10.10) и (10.6), получим наряду с обычным членом, пропорциональным $f_1(\boldsymbol{k})$, дополнительное слагаемое за счет того, что $N_q' \neq 0$, которое обусловливает эффект фононного увлечения. Для сохранения во всех членах одного порядка малости мы можем в члене фононного увлечения положить функции распределения электронов равными их равновесным значениям.

Таким образом, столкновительный член, связанный с фононным увлечением,

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}}^{\Phi} = \sum_{(q)} w(q)\, N_q' \{[f_0(\varepsilon_{k+q}) - f_0(\varepsilon_k)]\,\delta(\varepsilon_{k+q} - \varepsilon_k - \hbar\omega_q) - [f_0(\varepsilon_{k+q}) - f_0(\varepsilon_k)]\,\delta(\varepsilon_{k+q} - \varepsilon_k + \hbar\omega_q)\}, \tag{10.11}$$

где было использовано (10.6а). Если в этом выражении пренебречь неупругостью рассеяния, т. е. отбросить слагаемые $\hbar\omega_q$

в аргументах δ-функций, то все выражение будет равно нулю. В этом смысле в эффекте увлечения необходимо учитывать неупругость при поглощении и испускании фононов.

Так как изменение энергии электрона при поглощении или испускания фонона мало, то

$$f_0(\varepsilon_{k+q}) - f_0(\varepsilon_k) = \frac{\partial f_0}{\partial \varepsilon}(\varepsilon_{k+q} - \varepsilon_k), \tag{10.12}$$

где разность $\varepsilon_{k+q} - \varepsilon_k$ для первой квадратной скобки в (10.11) равна $\hbar w_q$, а для второй равна $-\hbar\omega_q$.

В результате

$$\left(\frac{\partial f}{\partial t}\right)^{\Phi}_{\text{ст}} = \sum_{(q)} 2w(q) N'_q \frac{\partial f_0}{\partial \varepsilon} \hbar\omega_q \delta(\varepsilon_{k+q} - \varepsilon_k), \tag{10.13}$$

где последовательно под знаком суммы, величина которой порядка $\hbar\omega_q = \hbar v_0 q$, не учитывается неупругость в аргументе δ-функции.

Выбирая в $\boldsymbol{q}$-пространстве полярную ось вдоль $\boldsymbol{k}$ и переходя от суммирования по $\boldsymbol{q}$ к интегрированию, получим

$$\left(\frac{\partial f}{\partial t}\right)^{\Phi}_{\text{ст}} = \frac{V}{(2\pi)^3} \int_{q_{\min}}^{q_{\max}} q^2\, dq \int_0^{\pi} \sin\vartheta\, d\vartheta \times$$

$$\times \int_0^{2\pi} d\varphi \left\{ 2w(q) N'_q \frac{\partial f_0}{\partial \varepsilon} \hbar v_0 q \delta(\varepsilon_{k+q} - \varepsilon_k) \right\}, \tag{10.14}$$

где V — объем основной области кристалла.

Подставляя сюда вместо $w(q)$ и N'_q (VIII.3.11в) и (10.8) и проводя интегрирование совершенно аналогично (VIII, § 4, п. 1), получим

$$\left(\frac{\partial f}{\partial t}\right)^{\Phi}_{\text{ст}} = \frac{C^2 \Omega_0 k_0 m^*}{9\pi M (\hbar k)^3} \frac{\partial f_0}{\partial \varepsilon} (\hbar \boldsymbol{k} \nabla T) \int_0^{2k} \tau_{\Phi}(q) q^3\, dq, \tag{10.15}$$

где $\Omega_0 = V/N$ — объем кристаллической ячейки.

Выражая C^2 из (VIII.4.11) через время релаксации τ', связанное с взаимодействием электрона с акустическими колебаниями, и вводя усреднение вида

$$\tilde{\tau}_{\Phi} = \frac{1}{4k^4} \int_0^{2k} \tau_{\Phi}(q) q^3\, dq, \tag{10.16}$$

получим

$$\left(\frac{\partial f}{\partial t}\right)^{\Phi}_{\text{ст}} = \frac{v_0^2}{T} \frac{\tilde{\tau}_{\Phi}}{\tau'} (\hbar \boldsymbol{k} \nabla T) \frac{\partial f_0}{\partial \varepsilon}. \tag{10.17}$$

В случае (10.9)

$$\tilde{\tau}_\Phi = \frac{1}{2A_l T^3 k^2} = \frac{\hbar^2}{4A_l m^* T^3 \varepsilon}, \tag{10.18}$$

где $\varepsilon = \hbar^2 k^2/2m^*$ — энергия электрона.

При учете фононного увлечения кинетическое уравнение в приближении скалярной эффективной массы имеет вид

$$\frac{\hbar \boldsymbol{k}}{m^*} \nabla_r f + \frac{1}{\hbar} \boldsymbol{F} \nabla_k f = \left(\frac{\partial f}{\partial t}\right)_{\text{ст}} + \left(\frac{\partial f}{\partial t}\right)^{\Phi}_{\text{ст}}, \tag{10.19}$$

где $\hbar \boldsymbol{k}/m^* = \boldsymbol{v}$ — скорость электрона, $\boldsymbol{F}$ — действующая на него сила и $(\partial f/\partial t)_{\text{ст}}$ — обычный член столкновения, связанный с неравновесностью функции распределения электронов f.

В левую часть (10.19) (за исключением слагаемого, пропорционального магнитному полю) можно подставить равновесную функцию распределения, тогда

$$\nabla_r f = \nabla_r f_0 = \frac{\partial f_0}{\partial T} \nabla T = -\frac{\partial f_0}{\partial \varepsilon} \left[\frac{\varepsilon - \zeta}{T} + \frac{\partial \zeta}{\partial T}\right] \nabla T \tag{10.20}$$

аналогично (2.5).

Подставляя (10.17) и (10.20) в (10.19), получим

$$-\frac{\partial f_0}{\partial \varepsilon}\left[\frac{\varepsilon + m^* v_0^2 (\tilde{\tau}_\Phi/\tau') - \zeta}{T} + \frac{\partial \zeta}{\partial T}\right]\left(\frac{\hbar \boldsymbol{k}}{m^*} \nabla T\right) + \frac{1}{\hbar} \boldsymbol{F} \nabla_k f = \left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -\frac{f - f_0}{\tau}, \tag{10.21}$$

где τ — полное время релаксации электронов, которое может определяться не только взаимодействием с акустическими колебаниями, но и другими механизмами рассеяния, например рассеянием на ионах примеси и т. д.

Из (10.21) видно, что при сделанных нами предположениях, учет эффекта фононного увлечения электронов сводится к появлению в кинетическом уравнении в члене, пропорциональном ∇T, дополнительного слагаемого в энергии ε, равного $m^* v_0^2 (\tilde{\tau}_\Phi/\tau')$, зависящего от энергии ε.

Это позволяет очень просто определять в кинетических явлениях аддитивные члены, соответствующие фононному увлечению.

3. Определим вклад, вносимый фононным увлечением в термоэдс и эффект Нернста.

Из (4.2) видно, что фононное увлечение определяет дополнительный вектор

$$\boldsymbol{\chi}^*_\Phi = \frac{k_0}{e} \tau \frac{m^* v_0^2}{k_0 T} \frac{\tilde{\tau}_\Phi}{\tau'} \nabla T, \tag{9.22}$$

который, согласно (2.21), приводит к дополнительному току

$$j_\Phi = \frac{ne^2}{m^*}\langle \chi_\Phi^* \rangle = \frac{ne^2}{m^*}\left(\frac{k_0}{e}\right)\frac{m^* v_0^2}{k_0 T}\left\langle \frac{\tau\tilde{\tau}_\Phi}{\tau'}\right\rangle \nabla T. \qquad (9.23)$$

Отсюда и из (4.5) для фононной части термоэдс следует

$$\alpha_\Phi = \left(\frac{k_0}{e}\right)\frac{m^* v_0^2}{k_0 T}\frac{\langle \tau\tilde{\tau}_\Phi/\tau'\rangle}{\langle\tau\rangle}. \qquad (10.24)$$

Если единственный существенный механизм рассеяния электронов — акустические колебания, то $\tau = \tau'$ и

$$\alpha_\Phi = \left(\frac{k_0}{e}\right)\frac{m^* v_0^2}{k_0 T}\frac{\langle\tilde{\tau}_\Phi\rangle}{\langle\tau\rangle}. \qquad (10.24а)$$

Из (10.18) следует, что $\langle\tilde{\tau}_\Phi\rangle \sim T^{-4}$, и так как $\langle\tau\rangle \sim T^{-3/2}$, то

$$\alpha_\Phi \sim T^{-7/2}, \qquad (10.24б)$$

т. е. быстро возрастает при понижении температуры.

Аналогично можно исследовать α_Φ при других механизмах рассеяния, когда $\tau \neq \tau'$.

Для определения эффекта фононного увлечения в явлении Нернста, исследуем структуру постоянной Q (6.4а). Рассмотрим для определенности случай слабого магнитного поля, когда в линейном приближении по H коэффициенты a_i и b_i равны (5.8) и (6.5).

Согласно предыдущему добавка $m^* v_0^2(\tilde{\tau}_\Phi/\tau')$, связанная с фононным увлечением, входит аддитивно в энергию ε в b_1 и b_2, так как в (6.3) эти коэффициенты стоят перед $\nabla_x T$ ($\nabla_y T = 0$!). Таким образом, в линейном приближении по H

$$Q_\Phi = \frac{a_1 b_2^\Phi - a_2 b_1^\Phi}{a_1^2 H} = \left(\frac{k_0}{e}\right)\frac{m^* v_0^2}{k_0 T}\frac{e}{m^* c}\frac{\langle\tau\rangle\left\langle\frac{\tilde{\tau}_\Phi}{\tau'}\tau^2\right\rangle - \langle\tau^2\rangle\left\langle\frac{\tilde{\tau}_\Phi}{\tau'}\tau\right\rangle}{\langle\tau\rangle^2}. \qquad (10.25)$$

Если электроны рассеиваются только на продольных акустических колебаниях, то $\tau = \tau'$ и

$$Q_\Phi = \left(\frac{k_0}{e}\right)\frac{m^* v_0^2}{k_0 T}\frac{e}{m^* c}\frac{\langle\tau\rangle\langle\tilde{\tau}_\Phi\tau\rangle - \langle\tau^2\rangle\langle\tilde{\tau}_\Phi\rangle}{\langle\tau\rangle^2}. \qquad (10.25а)$$

Если $\tilde{\tau}_\Phi$ определяется выражением (10.18), так что $\langle\tilde{\tau}_\Phi\rangle \sim T^{-4}$ и $\langle\tau\rangle \sim T^{-3/2}$, то

$$Q_\Phi \sim T^{-5}, \qquad (10.25б)$$

т. е. быстро возрастает при понижении температуры. Любопытно отметить, что если $\tau_\Phi \sim 1/q$ (а не $1/q^2$), то $Q_\Phi = 0$.

Из (10.24а), (4.5), (10.25а) и (6.6) следует, что по порядку величины

$$\frac{Q_\Phi/Q}{\alpha_\Phi/\alpha} \approx \frac{\langle \tau x \rangle}{\langle \tau \rangle} \frac{\zeta}{k_0 T}. \tag{10.26}$$

Для невырожденных полупроводников $(-\zeta/k_0T) \gg 1$, поэтому для них относительная роль эффекта увлечения в явлении Нернста больше, чем в термоэдс.

Эффект фононного увлечения в термомагнитных явлениях наблюдался впервые экспериментально в дырочном германии И. В. Мочан, Ю. Н. Образцовым и Т. В. Крыловой [1]) и независимо Херрингом и Джебл [2]).

Мы изложили только основы теории фононного увлечения электронов в полупроводнике. Наблюдаемые на опыте явления осложнены рядом обстоятельств.

Во-первых, при понижении температуры длина свободного пробега фонона $l_\Phi = v_0 \tau_\Phi$ (см. (10.18)) быстро возрастает и становится порядка линейных размеров L исследуемого образца, тогда $\tau_\Phi = L/v_0$ и не зависит от T и q. Очевидно, что при этом изменятся температурные зависимости α_Φ и Q_Φ, полученные выше.

Во-вторых, мы предполагали, что можно пренебречь обратным действием свободных электронов на неравновесную из-за градиента температуры функцию распределения фононов. Если фононная теплопроводность заметно снижается из-за рассеяния фононов на свободных электронах, то неравновесность фононной функции распределения уменьшается, и эффект фононного увлечения падает. Это явление наблюдается как при повышении концентрации носителей тока, так и при уменьшении температуры. Поэтому при понижении температуры для явления фононного увлечения наблюдается эффект насыщения.

В-третьих, новые и интересные явления возникают при изучении эффекта фононного увлечения в полупроводниках со сложной зонной структурой, особенно типа p-германия, когда имеются одновременно тяжелые и легкие дырки [3]).

По всем этим вопросам следует обращаться к специальной литературе; в частности, к обзору Херринга, напечатанному в сборнике трудов международной конференции по полупроводникам и фосфорам в Гармиш—Партенкирхене в 1956 г. [4]).

Учет эффекта фононного увлечения в других неизотермических явлениях производится аналогично предыдущему.

---

[1]) Мочан И. В., Образцов Ю. Н., Крылова Т. В.—ЖТФ, 1957, т. 27, с. 242.

[2]) Herring C., Geballe T. H.—Bullet. of. the Amer. Phys. Soc., 1956, ser. II, v. 1, p. 117.

[3]) См. там же.

[4]) Hableiter und Phosphore./Herausgegeben von M. Schön, H. Welker.—Berlin, 1958, p. 184.

§ 11. Квантовая теория гальвано- и термомагнитных явлений в полупроводниках [1])

Рассмотрим вначале вопрос о пределах применимости кинетического уравнения, которым мы пользовались в настоящей главе для расчета кинетических коэффициентов.

Мы рассмотрим две причины, ограничивающие применимость кинетического уравнения.

Первая из них связана с процессом рассеяния электронов проводимости, вторая — с квантованием движения электрона в магнитном поле.

1. Из соотношения Гейзенберга следует, что энергия электрона проводимости определяется с точностью до $\Delta\varepsilon > h/\Delta t$, где Δt — время квазистационарного состояния электрона, которое по порядку величины равно времени релаксации τ. С другой стороны, для применимости кинетического уравнения необходимо, чтобы неопределенность в энергии электрона $\Delta\varepsilon$ была много меньше интервала энергии k_0T, на котором существенно меняется функция распределения $f(\boldsymbol{k})$, входящая в кинетическое уравнение. Таким образом,

$$h/\tau < \Delta\varepsilon \ll k_0T.$$

Отсюда вытекает следующий критерий применимости кинетического уравнения:

$$\tau \gg h/k_0T. \tag{11.1}$$

Этот же критерий может быть получен и из других соображений.

Применение кинетического уравнения существенно основано на рассуждениях п. 2 Приложения 20. Вероятность перехода из состояния i в состояние f пропорциональна времени t и определяется выражением (П.20.7); квадратная скобка в нем как функция ω_{fi} обладает острым максимумом, полуширина которого порядка $\Delta\omega \approx 2\pi/t$ или в энергетических единицах $\hbar\Delta\omega = \Delta\varepsilon \approx h/t$. Так как в пределе для достаточно большого интервала времени t квадратная скобка переходит в δ-функцию от разности энергий конечного и начального состояний системы, то можно сказать, что величина $\Delta\varepsilon \approx h/t$ определяет точность выполнения закона сохранения энергии. В этом смысле можно смотреть на соотношение $\Delta\varepsilon t \approx h$ как на соотношение неопределенностей Гейзенберга для энергии и времени.

---

[1]) Изложение в этом параграфе носит более конспективный характер, чем в других частях книги; для более подробного ознакомления с вопросом рекомендуем книгу: Аскеров Б. М. Кинетические эффекты в полупроводниках.— Л.: Наука, 1970, гл. VII, VIII; см. также Зырянов П. С., Клингер М. И. Квантовая теория явлений электронного переноса в кристаллических полупроводниках.— М.: Наука, 1976.

Так как время измерения t должно быть меньше времени релаксации τ, а $\Delta\varepsilon$, как мы видели, должно быть много меньше k_0T, то мы опять приходим к критерию применимости кинетического уравнения (11.1).

Для электрона проводимости с эффективной массой m средняя энергия по порядку величины равна $m\bar{v}^2 = k_0T$, где $\bar{v}$ — средняя скорость электрона; соответствующая ему длина волны де-Бройля $\lambda = h/m\bar{v}$, а длина свободного пробега $l = \bar{v}\tau$. Используя эти соотношения, (11.1) может быть записано в виде

$$l \gg \lambda. \tag{11.2}$$

Для наглядной интерпретации критерия (11.2) отметим, что чем больше длина свободного пробега l электрона проводимости, тем меньше его взаимодействие с рассеивателем, и чем меньше дебройлевская длина волны λ, тем больше скорость (энергия) электрона.

Поскольку подвижность электрона $\mu = (e/m)\langle\tau\rangle$ (3.3), критерий (11.1) может быть записан в виде

$$\mu \gg \frac{eh}{mk_0T}. \tag{11.3}$$

Полагая $m = 10^{-27}$ *г* и $T = 300°$ K получим, что при комнатной температуре подвижность

$$\mu \gg 40 \text{ } см^2/в \cdot сек. \tag{11.3а}$$

Если подвижность носителей тока при комнатной температуре меньше правой части неравенства (11.3а), то кинетическое уравнение для их описания непригодно.

Критерий (11.1) (или эквивалентные ему (11.2) и (11.3)) справедлив в случае невырожденных электронов. Можно показать, что если электроны вырождены, то для кинетических эффектов, имеющих место в нулевом приближении по химическому потенциалу ζ (например, для электропроводности (7.8)), критерий применимости кинетического уравнения имеет вид

$$\tau \gg h/\zeta. \tag{11.4}$$

Не вполне выяснен вопрос, как выглядит критерий применимости кинетического уравнения для вырожденных электронов в случае кинетических эффектов, осуществляющихся только в приближении k_0T/ζ (например для термоэдс (7.11)).

2. В § 5 было отмечено, что если безразмерный параметр $\gamma\tau H \approx \mu H/c \gg 1$, то магнитное поле называется *сильным*; при обратном неравенстве оно называется *слабым*. Кинетическое уравнение применимо как в случае слабого, так и сильного поля.

Квантовомеханическое рассмотрение движения свободного электрона в магнитном поле показывает, что энергия, соответствующая движению в плоскости, перпендикулярной магнитному полю, квантована (VI.5.19б). Разность соседних уровней энергии равна $2\mu^* H = \hbar\omega_c$, где $\mu^* = e\hbar/2mc$ — эффективный магнетон Бора, а $\omega_c = eH/mc$ — циклотронная частота. Очевидно, что если

$$k_0 T \ll 2\mu^* H = \hbar\omega_c, \tag{11.5}$$

то пренебречь дискретностью уровней энергии нельзя, так как в этом случае тепловое движение только слегка расширяет квантовые уровни электрона.

Магнитные поля, удовлетворяющие неравенству (11.5), называются *квантующими*.

Отношение предельных (т. е. соответствующих равенству в (5.1)) сильных магнитных полей $H_{\text{сил}}$ к предельным квантующим магнитным полям $H_{\text{кв}}$ (11.5) равно

$$\frac{H_{\text{сил}}}{H_{\text{кв}}} = \frac{c}{\mu}\frac{2\mu^*}{k_0 T} = \frac{1}{2\pi}\frac{eh}{\mu m k_0 T} \ll 1. \tag{11.6}$$

Если выполняется (11.3), то $H_{\text{кв}} > H_{\text{сил}}$, и в этом случае существует область сильных магнитных полей, которые не являются еще квантующими, т. е. когда можно пользоваться кинетическим уравнением. В квантующих магнитных полях движение электрона не может быть описано посредством непрерывных значений энергии, импульса, координаты, поэтому кинетическое уравнение неприменимо.

3. Рассмотрим изотропный электронный проводник в изотермических условиях ($\nabla T = 0$), помещенный в магнитное поле $\boldsymbol{H}$, направленное вдоль оси z, и электрическое поле $\boldsymbol{E}$, лежащее в плоскости xy.

Феноменологические уравнения (1.1а) приобретают вид

$$j_x = \sigma_{xx} E_x - \sigma_{yx} E_y, \quad j_y = \sigma_{yx} E_x + \sigma_{xx} E_y, \tag{11.7}$$

так как $\sigma_{xx} = \sigma_{yy}$ и $\sigma_{xy} = -\sigma_{yx}$. Компоненты тензоров σ_{xx} и σ_{yx} совпадают с a_1 и a_2 в (5.4). В сильном магнитном поле, как следует из (5.15),

$$\sigma_{yx} = \frac{ne^2}{m}\frac{1}{\gamma H} = \frac{ecn}{H}. \tag{11.8}$$

Это значение σ_{yx}, не зависящее от рассеяния, получается и в квантовой механике.

Из (5.15) следует, что в случае сильного магнитного поля

$$\frac{\sigma_{xx}}{\sigma_{yx}} = \frac{a_1}{a_2} = \left\langle \frac{1}{\tau} \right\rangle \frac{mc}{eH} \approx \frac{1}{\gamma\tau H} \ll 1, \tag{11.9}$$

т. е. диссипативная составляющая тензора электропроводности σ_{xx} много меньше недиссипативной составляющей σ_{yx}. Соотно-

шение (11.9) сохраняется и для квантующих магнитных полей. На опыте обычно измеряют поперечное магнетосопротивление ρ_H, которое, как следует из (5.6), равно

$$\rho_H = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{yx}^2} \approx \frac{\sigma_{xx}}{\sigma_{yx}^2}. \qquad (11.10)$$

Диссипативная компонента тензора электропроводности σ_{xx} в квантующих магнитных полях была впервые вычислена румынским физиком Щ. Титейкой (1935). Теория Титейка основана на некоторых наглядных полуклассических представлениях, однако впоследствии она была строго подтверждена рассмотрением уравнений движения матрицы плотности (Е. Н. Адамс, Т. Д. Холштейн, 1959).

Титейка исходит из квантовомеханических уравнений движения электрона в скрещенных электрическом и магнитном полях (гл. VII, § 7, п. 2). При простой зоне энергия электрона (VII.7.23) в линейном приближении по электрическому полю E равна

$$\varepsilon = \varepsilon(0) + (2N+1)\mu^* H + \frac{\hbar^2 k_z^2}{2m} - \lambda^2 eEk_y, \qquad (11.11)$$

где у соответствующих величин опущен индекс c. Из (VII.7.21б) следует, что в том же приближении по электрическому полю энергия

$$\varepsilon = \varepsilon(0) + (2N+1)\mu^* H + \frac{\hbar^2 k_z^2}{2m} - eEx_0,$$

где $x_0 \equiv x_c$ — координата положения равновесия магнитного осциллятора. Таким образом, мы можем считать, что энергия электрона ε зависит от квантовых чисел N, k_z и x_0 и от электрического поля E, т. е. $\varepsilon(N, k_z, x_0; E) \equiv \varepsilon_\nu^E$, где $\nu \equiv \{N, k_z, x_0\}$. Считая, что в состоянии ν электрон «находится» в точке x_0, положим ток

$$j_x = -e \sum_{N,\, N',\, k_z,\, k_z'} \sum_{\substack{x_0 < 0 \\ x_0' > 0}} \{f_0(\varepsilon_\nu^E)[1 - f_0(\varepsilon_{\nu'}^E)] W_{\nu\nu'}^E - f_0(\varepsilon_{\nu'}^E)[1 - f_0(\varepsilon_\nu^E)] W_{\nu'\nu}^E\}. \qquad (11.12)$$

Здесь $f_0(\varepsilon_\nu^E)$ — равновесная функция распределения для электронов с энергией ε_ν^E, а $W_{\nu\nu'}^E$ — вероятность перехода электрона в единицу времени в результате рассеяния из состояния ν (с $x_0 < 0$) в состояние ν' (с $x_0' > 0$), т. е. вероятность такого рассеяния электрона, при котором он пересекает плоскость $x = 0$ слева направо; множитель $[1 - f_0(\varepsilon_{\nu'}^E)]$ в соответствии с принципом Паули учитывает вероятность того, что состояние ν' не занято другим электроном. Очевидно, что второе слагаемое в фигурной скобке определяет поток электронов через плоскость $x = 0$ справа налево.

Разлагая фигурную скобку в (11.12) по степеням электрического поля и ограничиваясь первым слагаемым, пропорциональным E, получим, сравнивая с (11.7)

$$\sigma_{xx}=e^2\sum_{\nu\nu'}\left[-\frac{\partial f_0(\varepsilon_\nu)}{\partial\varepsilon_\nu}\right]\frac{(x_0'-x_0)^2}{2}W_{\nu\nu'}, \qquad (11.13)$$

где все величины под знаком суммы взяты для поля $E=0$. Адамс и Холштейн, используя (11.13), определили для различных механизмов рассеяния в квантовом пределе $N=N'=0$ ($k_0T\ll\hbar\omega_c$) зависимость ρ_H (11.10) от магнитного поля H и температуры T; эта зависимость различна для невырожденных и вырожденных полупроводников.

В случае сильного вырождения носителей тока $-\partial f_0/\partial\varepsilon=\delta(\zeta-\varepsilon)$, где ζ—их химический потенциал. Из (VI.5.23а) следует, что в этом случае плотность числа состояний, входящая в (11.13), равна

$$g(\varepsilon)\sim\sum_N\frac{1}{\sqrt{\varepsilon-(2N+1)\mu^*H}}\delta(\zeta-\varepsilon)=\sum_N\frac{1}{\sqrt{\zeta-(2N+1)\mu^*H}}. \qquad (11.14)$$

Таким образом, при изменении магнитного поля, каждый раз, когда уровень Ландау пересекает уровень химического потенциала, $g(\varepsilon)$, а следовательно и σ_{xx}, резко возрастают. Эти осцилляции σ_{xx}, а следовательно и ρ_H (11.10), имеют период, определяемый из условия

$$2\Delta N\mu^*H+(2N+1)\mu^*\Delta H=0,$$

при $\Delta N=1$; отсюда

$$-\frac{\Delta H}{H^2}=\Delta\left(\frac{1}{H}\right)=\frac{2\mu^*}{\zeta}=\frac{e\hbar}{cm\zeta}. \qquad (11.15)$$

Эти осцилляции получили название *осцилляций Шубникова—де Гааза* по имени исследователей, открывших их на висмуте в 1930 г.; они наблюдаются только при сильном вырождении носителей тока, при этом их период $\Delta(1/H)$ зависит от концентрации электронов проводимости (поскольку ζ зависит от нее).

4. В 1964 г. В. Л. Гуревич и Ю. А. Фирсов теоретически предсказали новый вид осцилляций ρ_H, получивших название *магнетофононных*. Эти осцилляции связаны с неупругим рассеянием электронов на оптических фононах.

Рассмотрим магнетофононный резонанс с качественной точки зрения. В отсутствие рассеяния электроны проводимости, помещенные в скрещенные электрическое ($\boldsymbol{E}\parallel x$) и магнитное ($\boldsymbol{H}\parallel z$) поля испытывают в плоскости xy осцилляционное движение и дрейфуют с постоянной скоростью $v_y=c(E/H)$ в направлении оси y (см. (11.8)). Ток j_x создается только за счет рассеяния электронов. Вклад в ток j_x за счет электронов с энер-

гией ε и магнитным квантовым числом N будет пропорционален их числу $f_0(\varepsilon)\,g(\varepsilon, N)$, где $f_0(\varepsilon)$ — функция распределения Ферми, а $g(\varepsilon, N)$ — плотность числа состояний при фиксированном N, см. (VI.5.23а); очевидно, j_x будет также пропорционален вероятности перехода $(\varepsilon, N) \to (\varepsilon', N')$, равной

$$W_{NN'}(\varepsilon, \varepsilon')\,\delta(\varepsilon + \hbar\omega_0 - \varepsilon')\,[1 - f_0(\varepsilon')]\,g(\varepsilon', N'),$$

где $W_{NN'}(\varepsilon, \varepsilon')$ — некоторая плавная функция своих индексов и аргументов, δ-функция выражает закон сохранения энергии, связанный с поглощением[1]) оптического фонона $\hbar\omega_0$, а квадратная скобка определяет число свободных мест состояния с энергией ε'.

Полный ток j_x будет пропорционален произведению всех этих множителей, просуммированному по N и N' и проинтегрированному по ε и ε'. Выполняя интегрирование по ε' посредством δ-функции и используя для плотности состояний выражение (VI.5.23а), получим

$$j_x \sim \sum_{N,\,N'} \int d\varepsilon \frac{W_{NN'}(\varepsilon,\ \varepsilon + \hbar\omega_0)\, f_0(\varepsilon)\,[1 - f_0(\varepsilon + \hbar\omega_0)]}{\sqrt{\varepsilon - (N + 1/2)\,\hbar\omega_c}\,\sqrt{\varepsilon - [(N' + 1/2)\,\hbar\omega_c - \hbar\omega]}}, \qquad (11.16)$$

где мы положили $2\mu^* H = \hbar\omega_c$ ($\omega_c = eH/mc$ — циклотронная частота). Если величины, вычитаемые из ε под знаками корней в (11.16), различны, то интеграл по ε не содержит особенностей. Если же при изменении магнитного поля эти вычитаемые становятся равными, то интеграл логарифмически расходится, что означает резкое возрастание j_x.

Вычитаемые под знаками корней в (11.16) совпадают, если

$$(N' - N)\,\omega_c = \omega_0, \qquad (11.17)$$

т. е. ω_0 кратно ω_c. Учитывая, что $\omega_c = eH/mc$, получим для периода осцилляций j_x аналогично (11.15)

$$\Delta\left(\frac{1}{H}\right) = \frac{e}{mc\omega_0}. \qquad (11.18)$$

Существенное отличие магнетофононного резонанса от осцилляций Шубникова — де Гааза заключается в том, что последние наблюдаются только при вырождении носителей тока, и период осцилляций (11.15) зависит от их концентрации. Магнетофононный резонанс может наблюдаться как в вырожденных, так и в невырожденных полупроводниках, и период осцилляций (11.18) не зависит от концентрации. Из (11.18) может быть определена эффективная масса m, если известна предельная частота опти-

[1]) Результат не зависит от того, рассматриваем ли мы поглощение или испускание фонона.

ческих фононов ω_0. Опыты хорошо подтверждают все выводы теории магнетофононного резонанса [1]).

5. Рассмотрим кратко теорию термомагнитных явлений в квантующих магнитных полях.

При наличии градиента температуры $\nabla_x T$ вдоль оси x феноменологические уравнения (1.1а) имеют вид

$$\begin{aligned} j_x &= \sigma_{xx}\nabla_x\left(\frac{\zeta}{e}-\varphi\right)+\sigma_{xy}\nabla_y\left(\frac{\zeta}{e}-\varphi\right)-\beta_{xx}\nabla_x T, \\ j_y &= \sigma_{yx}\nabla_x\left(\frac{\zeta}{e}-\varphi\right)+\sigma_{yy}\nabla_y\left(\frac{\zeta}{e}-\varphi\right)-\beta_{yx}\nabla_x T. \end{aligned} \tag{11.19}$$

Так как химический потенциал ζ зависит от температуры, градиент которой направлен вдоль оси x, то $\nabla_y\left(\frac{\zeta}{e}-\varphi\right)=E_y$, $\nabla_x\left(\frac{\zeta}{e}-\varphi\right)\equiv E_x^*\neq E_x$. Для изотропного проводника $\sigma_{xx}=\sigma_{yy}$ и $\sigma_{xy}=-\sigma_{yx}$. В нулевом приближении по рассеянию только σ_{yx} и β_{yx} отличны от нуля; кроме того, в дополнение к (11.9) в сильных полях ($\gamma\tau H\gg 1$), как следует из (6.8),

$$\frac{\beta_{xx}}{\beta_{yx}}=\frac{b_1}{b_2}\approx\frac{1}{\gamma\tau H}\ll 1. \tag{11.20}$$

В результате для термоэдс и постоянной Нернста в сильных магнитных полях в наинизшем приближении по рассеянию получим

$$\alpha_H=\frac{|E_x^*|}{|\nabla_x T|}=\frac{\beta_{yx}}{\sigma_{yx}}, \tag{11.21}$$

$$Q=-\frac{E_y}{\nabla_x T H}=\frac{\sigma_{yx}\beta_{xx}-\sigma_{xx}\beta_{yx}}{\sigma_{yx}^2 H}, \tag{11.21а}$$

как это следует из (6.11) и (6.4а).

Для определения α_H и Q надо в дополнение к (11.8) и (11.13) вычислить β_{yx} и β_{xx}. При этом возникает следующая принципиальная трудность. В классическом случае при наличии градиента температуры надо рассматривать температуру T как функцию координаты x, т. е. $T=T(x)$. В квантовом случае это предположение теряет смысл, так как координата электрона x в квантующем магнитном поле не является квантовым числом, т. е. не имеет определенного значения. Исходя из принципа локального термодинамического равновесия, А. И. Ансельм и Б. М. Аскеров (1960) предположили, что при наличии градиента температуры вдоль x температура является функцией положения равновесия магнитного осциллятора x_0 (VI.5.18а). Рассуждая

[1]) Для подробного ознакомления с вопросом см. обзор Парфеньев Р. В., Харус Г. И., Цидильковский И. М., Шалыт С. С. Магнетофононный резонанс в полупроводниках.— УФН, 1974, т. 112, с. 3—36.

так же, как при выводе σ_{xx} (11.13), можно показать, что

$$\beta_{xx}=\frac{e}{2T}\sum_{\nu\nu'}\left[\frac{\partial f_0(\varepsilon_\nu)}{\partial\varepsilon_\nu}\right](\varepsilon_\nu-\zeta)(x_0-x_0')^2 W_{\nu\nu'}. \qquad (11.22)$$

Так как нас интересуют эффекты, линейные по электрическому полю E и градиенту температуры $\nabla_x T$, то можно в (11.22) считать все величины для $E=0$. Заметим, что это выражение было обосновано в работе А. И. Ансельма, Ю. Н. Образцова и Р. Г. Тарханяна (1965), аналогично тому, как было обосновано (11.13) в работе Адамса и Холштейна.

Если исходить из квантовомеханического выражения для плотности тока $j_y^{(\nu)}$ для электрона в состоянии $\nu=\{N, x_0, k_z\}$ и предположения $T=T(x_0)$, то плотность тока

$$j_y=\sum_\nu j_y^{(\nu)} f_0\left(\frac{\varepsilon_\nu-\zeta(x_0)}{k_0T(x_0)}\right). \qquad (11.23)$$

Разлагая функцию Ферми f_0 по степеням приращения Δx_0 и отождествляя $\Delta T/\Delta x_0=\nabla_x T$, получим для

$$\beta_{yx}=\frac{(2m)^{1/2}e\omega_c}{4\pi^2\hbar}\sum_N\left(N+\frac{1}{2}\right)\int_{\varepsilon_N}^{\infty}\left(\frac{\varepsilon-\zeta}{T}\right)\frac{\partial f_0}{\partial\varepsilon}\frac{d\varepsilon}{\sqrt{\varepsilon-\varepsilon_N}}, \qquad (11.23а)$$

где $\varepsilon_N=(N+1/2)\hbar\omega_c$. Оказалось, однако, что это выражение для β_{yx} не удовлетворяет принципу Онсагера (1.5а). Правильное выражение для β_{yx} было получено Ю. Н. Образцовым (1963). Он обратил внимание на то, что в случае квантующих магнитных полей при вычислении недиссипативного тока j_y следует к объемному току, определяемому формулой (11.23а), прибавить поверхностные токи, текущие по граням образца перпендикулярным оси x. В отсутствие градиента температуры ($\nabla_x T=0$) эти токи взаимно компенсируются, обусловливая лишь диамагнетизм электронов проводимости (гл. VI, § 5). В случае же, когда $\nabla_x T\neq 0$, такая компенсация не имеет места и возникает ток вдоль оси y, пропорциональный $\nabla_x T$. Учет этого тока приводит к следующему простому выражению для β_{yx}:

$$\beta_{yx}=-cS/H, \qquad (11.24)$$

где S — энтропия электронов в 1 $см^3$, c — скорость света. Отсюда из (11.21) и (11.8) получим

$$\alpha_H=s/c, \qquad (11.25)$$

где $s=S/n$ — энтропия в расчете на 1 электрон.

Как показал Образцов, в случае невырожденных электронов

$$\alpha_H=\frac{k_0}{e}\left[\frac{3}{2}+\frac{\hbar\omega_c}{2k_0T}\operatorname{cth}\left(\frac{\hbar\omega_c}{2k_0T}\right)-\frac{\zeta(H)}{k_0T}\right]. \qquad (11.25а)$$

В квазиклассическом случае: $(\hbar\omega_c/2k_0T) \ll 1$; тогда $(\hbar\omega_c/2k_0T) \times \times \operatorname{cth}(\hbar\omega_c/2k_0T) \approx 1$ и

$$\alpha_H = \frac{k_0}{e}\left[\frac{5}{2} - \frac{\zeta}{k_0T}\right],$$

что совпадает с (IX.6.16).

Формула (11.25а) получила подтверждение в опытах с InSb (И. Л. Дричко и И. В. Мочан, 1964; С. Пури и Г. Джеблл, 1964) [1]).

Используя (11.8), (11.13), (11.22) и (11.24), можно определить константу Нернста Q (11.21а) для различных механизмов рассеяния и сравнить результаты с опытом. К сожалению, таких опытов, которые могли бы быть однозначно сопоставлены с теорией, не имеется.

---

[1]) Расхождение эксперимента с теорией при $(\hbar\omega_c/2k_0T) > 1$ не объясняется ни спиновым расщеплением уровней Ландау, ни непараболичностью зоны; природа этого расхождения остается невыясненной.

ПРИЛОЖЕНИЯ

Приложение 1

Рассмотрим косоугольные координаты на плоскости. Обобщение результатов на трехмерный случай не представляет затруднений.

На рис. П.1 изображены оси X_1 и X_2 прямоугольной системы координат и оси Ξ_1 и Ξ_2 косоугольной системы, имеющие общее начало O.

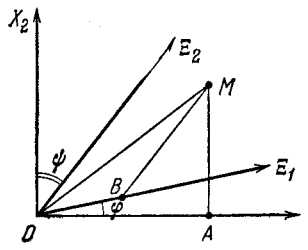


Рис. П. 1.

Положение некоторой точки M определяется радиусом-вектором $\overline{OM}$, прямоугольными координатами $\overline{OA}=x_1$ и $\overline{AM}=x_2$ или косоугольными координатами $\overline{OB}=\xi_1$ и $\overline{BM}=\xi_2$ $(\overline{BM}\parallel\Xi_2)$. Вектор $\overline{OM}$ можно рассматривать как сумму векторов $\overline{OB}$ и $\overline{BM}$, т. е.

$$\overline{OM}=\overline{OB}+\overline{BM}. \qquad (П.1.1)$$

Проектируя обе части этого равенства на оси X_1 и X_2, получим

$$x_1=\xi_1\cos\varphi+\xi_2\sin\psi, \quad x_2=\xi_1\sin\varphi+\xi_2\cos\psi.$$

Решая эту систему относительно ξ_1 и ξ_2, получим

$$\xi_1=\alpha_{11}x_1+\alpha_{12}x_2, \quad \xi_2=\alpha_{21}x_1+\alpha_{22}x_2, \qquad (П.1.2)$$

где

$$\alpha_{11}=\frac{\cos\psi}{\cos(\varphi+\psi)}, \quad \alpha_{12}=-\frac{\sin\psi}{\cos(\varphi+\psi)}, \quad \alpha_{21}=-\frac{\sin\varphi}{\cos(\varphi+\psi)},$$

$$\alpha_{22}=\frac{\cos\varphi}{\cos(\varphi+\psi)}. \qquad (П.1.3)$$

Аналогично в случае трехмерной косоугольной системы координат, начало которой совпадает с началом декартовой системы, координаты ξ_i являются однородными линейными функциями декартовых координат x_k

$$\xi_i=\sum_{k=1}^{3}\alpha_{ik}x_k \quad (i=1,2,3), \qquad (П.1.3а)$$

где α_{ik} зависят от углов, которые оси Ξ_i образуют с осями X_k.

Приложение 2

Для решения уравнения (3.7) представим неизвестный вектор разложенным по трем некомпланарным векторам $[a_1a_2]$, $[a_2a_3]$ и $[a_3a_1]$

$$\boldsymbol{b}=\alpha[a_1a_2]+\beta[a_2a_3]+\gamma[a_3a_1]. \qquad (П.2.1)$$

Такое разложение всегда возможно и единственно [1]). Здесь α, β и γ — скалярные множители, которые необходимо определить. Подставляя (П.2.1) в (3.7), получим

$$\boldsymbol{b}\boldsymbol{a}_1=\beta(\boldsymbol{a}_1[\boldsymbol{a}_2\boldsymbol{a}_3])=2\pi g_1,$$
$$\boldsymbol{b}\boldsymbol{a}_2=\gamma(\boldsymbol{a}_2[\boldsymbol{a}_3\boldsymbol{a}_1])=2\pi g_2,$$
$$\boldsymbol{b}\boldsymbol{a}_3=\alpha(\boldsymbol{a}_3[\boldsymbol{a}_1\boldsymbol{a}_2])=2\pi g_3,$$

откуда

$$\alpha=\frac{2\pi g_3}{\Omega_0},\quad \beta=\frac{2\pi g_1}{\Omega_0},\quad \gamma=\frac{2\pi g_2}{\Omega_0}. \tag{П.2.2}$$

Из (П.2.1) и (П.2.2) следует (3.8).

Приложение 3

1. Рассмотрим две прямоугольные координатные системы (x_1, x_2, x_3) и (x_1', x_2', x_3'), имеющие общее начало координат O. Косинусы углов между осями обеих систем обозначим через $\alpha_{ik}=\cos(\widehat{x_i x_k'})$, где i и k независимо принимают значения 1, 2 и 3. Известно, что

$$\sum_{i=1}^{3}\alpha_{ik}\alpha_{il}=\sum_{i=1}^{3}\alpha_{ki}\alpha_{li}=\delta_{kl}. \tag{П.3.1}$$

Для $k=l$ ($\delta_{kl}=1$) эти равенства означают, что сумма квадратов направляющих косинусов равна единице, а для $k\neq l$ $\delta_{kl}=0$ — условие перпендикулярности осей x_k и x_l (или x_k' и x_l'). Известно, что проекция на данное направление (например, x_1') геометрической суммы (радиуса-вектора $\boldsymbol{r}$) равна сумме проекций составляющих, поэтому

$$\left.\begin{aligned} x_1'&=\alpha_{11}x_1+\alpha_{12}x_2+\alpha_{13}x_3,\\ x_2'&=\alpha_{21}x_1+\alpha_{22}x_2+\alpha_{23}x_3,\\ x_3'&=\alpha_{31}x_1+\alpha_{32}x_2+\alpha_{33}x_3,\end{aligned}\right\} \tag{П.3.2}$$

и аналогично

или

$$x_i'=\sum_{k=1}^{3}\alpha_{ik}x_k. \tag{П.3.3}$$

Запишем (П.3.2) более компактно:

$$\boldsymbol{r}'=\boldsymbol{\alpha}\boldsymbol{r}. \tag{П.3.4}$$

Здесь

$$\boldsymbol{\alpha}=\begin{pmatrix}\alpha_{11} & \alpha_{12} & \alpha_{13}\\ \alpha_{21} & \alpha_{22} & \alpha_{23}\\ \alpha_{31} & \alpha_{32} & \alpha_{33}\end{pmatrix} \tag{П.3.5}$$

— *матрица* линейного преобразования (П.3.2). Матрица $\boldsymbol{\alpha}$ (П.3.5) *третьего ранга*, т. е. квадратная таблица, состоящая из трех строк, трех столбцов и $3^2=9$ элементов α_{ik}. Если элементы матрицы α_{ik} удовлетворяют условиям (П.3.1), то линейное преобразование (П.3.2) называется *ортогональным*. Обратное преобразование от штрихованных координат x_k' к нештрихованным x_i может быть

[1]) Смирнов В. И., 1954, т. 2, § 102.

получено аналогично (П.3.2)

$$\begin{aligned} x_1 &= \alpha_{11}x_1' + \alpha_{21}x_2' + \alpha_{31}x_3', \\ x_2 &= \alpha_{12}x_1' + \alpha_{22}x_2' + \alpha_{32}x_3', \\ x_3 &= \alpha_{13}x_1' + \alpha_{23}x_2' + \alpha_{33}x_3', \end{aligned} \tag{П.3.6}$$

что может быть записано подобно (П.3.4)

$$\boldsymbol{r} = \boldsymbol{\alpha}^{-1}\boldsymbol{r}', \tag{П.3.6а}$$

где $\boldsymbol{\alpha}^{-1}$ — матрица *обратного преобразования* (П.3.6).

Мы видим, что матрица обратного ортогонального преобразования $\boldsymbol{\alpha}^{-1}$ (П.3.6) может быть получена из матрицы прямого ортогонального преобразования $\boldsymbol{\alpha}$ (П.3.2) посредством замены строк на столбцы, т. е.

$$(\boldsymbol{\alpha}^{-1})_{ik} = (\boldsymbol{\alpha})_{ki}. \tag{П.3.7}$$

Скалярное произведение векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ равно

$$\boldsymbol{ab} = ab\cos(\widehat{\boldsymbol{a}, \boldsymbol{b}}) = \sum_{i=1}^{3} a_i b_i = \sum_{i=1}^{3} a_i' b_i', \tag{П.3.8}$$

где a_i', b_i' — составляющие векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ в координатной системе (x_1', x_2', x_3'). Последний переход в этих равенствах следует как из геометрического смысла скалярного произведения, так и формально, если выразить a_i и b_i через a_i' и b_i' по формулам (П.3.6) и воспользоваться условиями ортогональности (П.3.1).

2. Обобщая (П.3.2), запишем общее однородное линейное преобразование n переменных в виде

$$\begin{aligned} x_1' &= \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n, \\ x_2' &= \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n, \\ &\cdots\cdots\cdots\cdots\cdots \\ x_n' &= \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots + \alpha_{nn}x_n, \end{aligned} \tag{П.3.9}$$

или

$$x_i' = \sum_{k=1}^{n} \alpha_{ik}x_k \quad (i = 1, 2, \ldots, n), \tag{П.3.10}$$

что может быть записано аналогично (П.3.4)

$$\boldsymbol{r}' = \boldsymbol{\alpha}\boldsymbol{r}. \tag{П.3.11}$$

Здесь $\boldsymbol{r} \equiv \{x_1, x_2, \ldots, x_n\}$ и $\boldsymbol{r}' \equiv \{x_1', x_2', \ldots, x_n'\}$ — радиусы-векторы в n-мерном пространстве, а $\boldsymbol{\alpha}$ — матрица n-го ранга линейного преобразования (П.3.9), т. е.

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix}. \tag{П.3.12}$$

В общем случае, переменные x_i', x_k и элементы матрицы α_{ik} — комплексные числа.

Матрица n-го ранга (П.3.12) — квадратная таблица, состоящая из n строк, n столбцов и n^2 элементов α_{ik}. Про элементы α_{ii} говорят, что они расположены вдоль *главной диагонали* матрицы (т. е. диагонали, идущей от верхнего левого угла квадрата к его правому нижнему углу).

Последовательное применение двух линейных преобразований вида (П.3.11) с матрицами α и β дает

$$\left.\begin{aligned} \boldsymbol{r}' &= \alpha \boldsymbol{r}, \quad \boldsymbol{r}'' = \beta \boldsymbol{r}', \\ \text{откуда} \quad \boldsymbol{r}'' &= \beta \boldsymbol{r}' = \beta\alpha \boldsymbol{r}. \end{aligned}\right\} \tag{П.3.13}$$

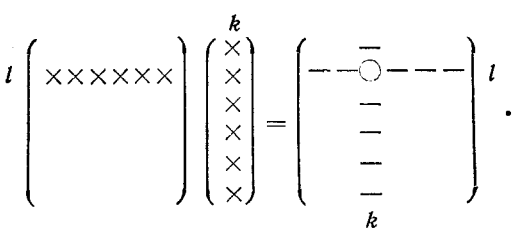
В раскрытом виде

$$x_l'' = \sum_{i=1}^{n} \beta_{li} x_i' = \sum_{i=1}^{n} \sum_{k=1}^{n} \beta_{li}\alpha_{ik} x_k = \sum_{k=1}^{n} (\beta\alpha)_{lk} x_k, \tag{П.3.14}$$

где

$$(\beta\alpha)_{lk} = \sum_{i=1}^{n} \beta_{li}\alpha_{ik}. \tag{П.3.15}$$

Последнее соотношение дает *правило умножения матриц*. Для того чтобы получить lk-й элемент матрицы, равной произведению матриц $\beta\alpha$, надо перемножить элементы l-й строки матрицы β на элементы k-го столбца матрицы α и сложить их. Это правило может быть схематически изображено так:



(П.3.16)

В ряде случаев целесообразно обобщить понятие матрицы (П.3.5) на прямоугольные таблицы, в которых число строк не равно числу столбцов. Правило умножения матриц (П.3.15) непосредственно обобщается на случай, когда число n столбцов в матрице β равно числу n строк в матрице α (в этом случае индекс i в (П.3.15) пробегает значения от 1 до n). Произведение $\beta\alpha$ есть в этом случае прямоугольная матрица, число строк в которой равно числу строк в β, а число столбцов — числу столбцов в α.

В важном случае умножения квадратной матрицы на матрицу из одного столбца мы получаем матрицу из одного столбца.

Очевидно, что произведение матриц, вообще говоря, *не коммутативно* (см. П.3.16), т. е.

$$\beta\alpha \neq \alpha\beta, \tag{П.3.17}$$

но в то же время удовлетворяет ассоциативному закону:

$$\gamma(\beta\alpha) = (\gamma\beta)\alpha, \tag{П.3.18}$$

как легко проверить, пользуясь правилом умножения матриц (П.3.15).

3. Если в квадратной матрице отличны от нуля только элементы, стоящие вдоль главной диагонали, она называется диагональной и ее элементы равны

$$D_{ik} = A_i \delta_{ik}, \tag{П.3.19}$$

где δ_{ik} — символ Кроникера.

Единичной матрицей $\boldsymbol{E}$ называется диагональная матрица, все элементы которой равны единице

$$E_{ik} = \delta_{ik}. \tag{П.3.20}$$

Пользуясь правилом умножения матриц (П.3.15), легко показать, что для любой матрицы α

$$\boldsymbol{E}\alpha = \alpha\boldsymbol{E} = \alpha. \tag{П.3.21}$$

Под *нулевой* матрицей $\mathbf{0}$ понимают матрицу, все элементы которой равны нулю. Очевидно, что для произвольной матрицы $\boldsymbol{\alpha}$

$$\mathbf{0}\boldsymbol{\alpha}=\boldsymbol{\alpha}\mathbf{0}=\mathbf{0}. \tag{П.3.22}$$

Под произведением матрицы $\boldsymbol{\alpha}$ на число c понимают матрицу с элементами $c\alpha_{ik}$. Под суммой матриц $\boldsymbol{\alpha}$ и $\boldsymbol{\beta}$ одного ранга понимают матрицу $\boldsymbol{\gamma}$ с элементами $\gamma_{ik}=\alpha_{ik}+\beta_{ik}$. Матрица $\boldsymbol{\alpha}$ равна матрице $\boldsymbol{\beta}$ тогда и только тогда, когда для всех элементов $\alpha_{ik}=\beta_{ik}$.

Шпуром (немецкое слово spur — след) или следом матрицы $\boldsymbol{\alpha}$ называется сумма ее диагональных элементов

$$\mathrm{Sp}\,\boldsymbol{\alpha}=\sum_i \alpha_{ii}. \tag{П.3.23}$$

Иногда шпур матрицы $\boldsymbol{\alpha}$ обозначается как $\mathrm{Tr}\,\boldsymbol{\alpha}$ (английское слово trace — след).

Шпур произведения матриц не зависит от их порядка; в самом деле,

$$\mathrm{Sp}\,(\boldsymbol{\beta}\boldsymbol{\alpha})=\sum_i(\boldsymbol{\beta}\boldsymbol{\alpha})_{ii}=\sum_i\sum_k\beta_{ik}\alpha_{ki}=\sum_k\Big(\sum_i\alpha_{ki}\beta_{ik}\Big)=\sum_k(\boldsymbol{\alpha}\boldsymbol{\beta})_{kk}=\mathrm{Sp}\,(\boldsymbol{\alpha}\boldsymbol{\beta}). \tag{П.3.24}$$

Матрица, *обратная* $\boldsymbol{\alpha}$, обозначаемая через $\boldsymbol{\alpha}^{-1}$, осуществляет преобразование, обратное преобразованию (П.3.11), так что

$$\boldsymbol{r}=\boldsymbol{\alpha}^{-1}\boldsymbol{r}'. \tag{П.3.25}$$

Подставляя сюда (П.3.11), получим

$$\boldsymbol{r}=\boldsymbol{\alpha}^{-1}\boldsymbol{\alpha}\boldsymbol{r}, \tag{П.3.26}$$

откуда

$$\left.\begin{aligned}&\boldsymbol{\alpha}^{-1}\boldsymbol{\alpha}=\boldsymbol{E}\text{—единичная матрица,}\\ \text{аналогично}\quad&\boldsymbol{\alpha}\boldsymbol{\alpha}^{-1}=\boldsymbol{E}.\end{aligned}\right\} \tag{П.3.27}$$

Эти соотношения справедливы, конечно, и для ортогональных преобразований (П.3.4) и (П.3.6а).

Мы видели (см. (П.3.7)), что матрица, обратная ортогональному преобразованию $\boldsymbol{\alpha}^{-1}$, есть матрица, получаемая из $\boldsymbol{\alpha}$ заменой в ней строк на столбцы (это эквивалентно «отражению» всех элементов матрицы в главной диагонали); такая матрица называется *транспонированной*. Обозначая транспонированную матрицу через $\tilde{\boldsymbol{\alpha}}$, т. е. $\tilde{\alpha}_{ik}=\alpha_{ki}$, имеем для ортогонального преобразования

$$\boldsymbol{\alpha}^{-1}=\tilde{\boldsymbol{\alpha}}, \tag{П.3.28}$$

как это следует из (П.3.7).

Если $\boldsymbol{\alpha}$ — матрица n-го ранга, то матричное уравнение (П.3.27) эквивалентно n^2 алгебраических уравнений первой степени для n^2 неизвестных α^{-1}_{ik}. Так как в n из этих уравнений правые части равны единице (в остальных они равны нулю), то n^2 неоднородных уравнений имеют решение для неизвестных α^{-1}_{ik} только тогда, когда определитель $|\alpha_{ik}|$, составленный из элементов матрицы $\boldsymbol{\alpha}$, отличен от нуля. Таким образом, матрица $\boldsymbol{\alpha}$ имеет обратную матрицу $\boldsymbol{\alpha}^{-1}$ только в том случае, когда $|\alpha_{ik}|\neq 0$, т. е. матрица $\boldsymbol{\alpha}$ *не сингулярная*.

Легко показать, что

$$(\boldsymbol{\beta}\boldsymbol{\alpha})^{-1}=\boldsymbol{\alpha}^{-1}\boldsymbol{\beta}^{-1} \tag{П.3.29}$$

и

$$\widetilde{(\boldsymbol{\beta}\boldsymbol{\alpha})}=\tilde{\boldsymbol{\alpha}}\tilde{\boldsymbol{\beta}}. \tag{П.3.30}$$

Из (П.3.27) и (П.3.28) следует, что для ортогонального преобразования

$$\tilde{\boldsymbol{\alpha}}\boldsymbol{\alpha}=\boldsymbol{E}. \tag{П.3.31}$$

Так как правило умножения матриц (П.3.15) совпадает с правилом умножения

определителей [1]), то

$$|\tilde{\alpha}||\alpha| = |E| = 1, \qquad (П.3.32)$$

где $|\alpha|$ — определитель матрицы α. Так как определитель не меняется при замене в нем строк столбцами [2]), то $|\tilde{\alpha}| = |\alpha|$ и, следовательно,

$$|\alpha|^2 = 1. \qquad (П.3.33)$$

Следовательно, при ортогональном преобразовании (П.3.2) определитель матрицы преобразования

$$|\alpha| = \pm 1. \qquad (П.3.33а)$$

Нетрудно показать, что $+1$ соответствует простому вращению системы, а -1 — вращению, сопровождаемому инверсией, т. е. переходу от правой координатной системы к левой (или наоборот). В самом деле, простой поворот можно представить себе как непрерывный процесс вращения из исходного состояния, в котором обе системы совпадают и которому соответствует тождественное преобразование с матрицей $E_{ik} = \delta_{ik}$. Так как этому преобразованию соответствует определитель $|E_{ik}| = |\delta_{ik}| = 1$ и при непрерывном вращении он не может измениться скачком, то простому вращению соответствует определитель, равный $+1$.

При инверсии J: $x_i' = -x_i$ $(i = 1, 2, 3)$ и этому преобразованию соответствует определитель $|J| = |-\delta_{ik}| = -1$. Отсюда сразу следует, что любому вращению, сопровождаемому инверсией, тоже соответствует определитель, равный -1. Ортогональные преобразования с определителем, равным $+1$, получили название *собственных вращений* (или просто вращений), а ортогональные преобразования с определителем, равным -1 — *несобственных вращений*.

4. Рассмотрим некоторые важные для приложений матрицы. Обозначим через α^* матрицу *комплексно-сопряженную* с α, так что $(\alpha^*)_{ik} = \alpha^*_{ik}$.

Под *сопряженной* матрицей α^+ понимается транспонированная, комплексно-сопряженная матрица, т. е. $\alpha^+ = \tilde{\alpha}^*$, так что $\alpha^+_{ik} = \alpha^*_{ki}$.

Самосопряженная или *эрмитова* матрица α обладает свойством $\alpha^+ = \alpha$; элементы эрмитовой матрицы, расположенные симметрично относительно главной диагонали, — комплексно-сопряженные.

Унитарной матрицей называется такая, для которой сопряженная равна обратной, т. е.

$$U^+ = U^{-1} \quad \text{или} \quad UU^+ = U^+U = E. \qquad (П.3.34)$$

Унитарная матрица осуществляет *унитарное преобразование*, являющееся обобщением ортогонального преобразования. Для того чтобы показать это, введем понятие о *внутреннем* или *эрмитовом скалярном произведении* двух n-мерных комплексных векторов $a\,(a_1, a_2, \ldots, a_n)$ и $b\,(b_1, b_2, \ldots, b_n)$, которое по определению равно

$$(ab) = \sum_{i=1}^{n} a_i^* b_i. \qquad (П.3.35)$$

Если a и b — трехмерные вещественные векторы, определение (П.3.35) переходит в обычное скалярное произведение (П.3.8). Покажем далее, что

$$(a, \alpha b) = (\alpha^+ a, b), \qquad (П.3.36)$$

где α — матрица линейного преобразования n-го ранга. В самом деле, левая часть (П.3.36) равна

$$\sum_i a_i^* (\alpha b)_i = \sum_{ik} a_i^* \alpha_{ik} b_k,$$

---

[1]) Смирнов В. И., т. 3, ч. I, § 1, п. 6.

[2]) См. там же.

а правая часть —

$$\sum_i (\alpha^+ a)_i^* b_i = \sum_{ik} a_k^* \alpha_{ki} b_i,$$

что отличается от левой части только обозначением немых индексов суммирования.

Применим теперь к векторам $\boldsymbol{a}$ и $\boldsymbol{b}$ в (П.3.35) унитарное преобразование $\boldsymbol{U}$, тогда

$$(\boldsymbol{Ua}, \boldsymbol{Ub}) = (\boldsymbol{U}^+\boldsymbol{Ua}, \boldsymbol{b}) = (\boldsymbol{Ea}, \boldsymbol{b}) = (\boldsymbol{a}, \boldsymbol{b}), \qquad (П.3.37)$$

где мы использовали (П.3.36) и (П.3.34).

Выражение (П.3.37) является естественным обобщением соотношения (П.3.8), инвариантного относительно ортогонального преобразования $\boldsymbol{\alpha}$.

В заключение этого пункта введем важное понятие о *преобразовании подобия* матрицы $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha}' = \boldsymbol{s}^{-1}\boldsymbol{\alpha}\boldsymbol{s}, \qquad (П.3.38)$$

где $\boldsymbol{s}$ — произвольная не сингулярная матрица (т. е. матрица, имеющая обратную).

Умножая (П.3.38) слева на $\boldsymbol{s}$ и справа на $\boldsymbol{s}^{-1}$ и учитывая, что $\boldsymbol{s}^{-1}\boldsymbol{s} = \boldsymbol{E}$, получим

$$\boldsymbol{\alpha} = \boldsymbol{s}\boldsymbol{\alpha}'\boldsymbol{s}^{-1}. \qquad (П.3.38а)$$

Можно показать, что любое матричное уравнение не меняется (инвариантно) при преобразовании подобия. Например, матричное уравнение

$$\boldsymbol{\beta}\boldsymbol{\alpha} + \boldsymbol{\gamma} = \boldsymbol{\varepsilon} \qquad (П.3.39)$$

при преобразовании подобия (П.3.38а) приобретает вид

$$\boldsymbol{s}\boldsymbol{\beta}'\boldsymbol{s}^{-1}\boldsymbol{s}\boldsymbol{\alpha}'\boldsymbol{s}^{-1} + \boldsymbol{s}\boldsymbol{\gamma}'\boldsymbol{s}^{-1} = \boldsymbol{s}\boldsymbol{\varepsilon}'\boldsymbol{s}^{-1}.$$

Умножая это выражение слева на $\boldsymbol{s}^{-1}$ и справа на $\boldsymbol{s}$, получим

$$\boldsymbol{\beta}'\boldsymbol{\alpha}' + \boldsymbol{\gamma}' = \boldsymbol{\varepsilon}', \qquad (П.3.39а)$$

т. е. уравнение (П.3.39) инвариантно относительно преобразования подобия.

Покажем еще, что шпур матрицы не меняется при преобразовании подобия

$$\mathrm{Sp}\,\boldsymbol{\alpha} = \mathrm{Sp}\,(\boldsymbol{s}\boldsymbol{\alpha}'\boldsymbol{s}^{-1}) = \mathrm{Sp}\,(\boldsymbol{s}^{-1}\boldsymbol{s}\boldsymbol{\alpha}') = \mathrm{Sp}\,(\boldsymbol{E}\boldsymbol{\alpha}') = \mathrm{Sp}\,\boldsymbol{\alpha}', \qquad (П.3.40)$$

где мы воспользовались (П.3.24).

5. Вектор $\boldsymbol{r}'$ в (П.3.11), вообще говоря, не пропорционален вектору $\boldsymbol{r}$. Если же для некоторого значения $\boldsymbol{r} = \boldsymbol{x}\,(x_1, x_2, \ldots, x_n)$

$$\boldsymbol{\alpha}\boldsymbol{x} = \lambda\boldsymbol{x}, \qquad (П.3.41)$$

где λ — скаляр, то $\boldsymbol{x}$ называется *собственным вектором*, а λ — *собственным значением* матрицы $\boldsymbol{\alpha}$.

В проекциях на оси уравнение (П.3.41) имеет вид

$$\left.\begin{aligned} &\sum_{k=1}^{n} \alpha_{ik} x_k = \lambda x_i, \\ \text{или} \quad &\\ &\sum_{k=1}^{n} (\alpha_{ik} - \lambda\delta_{ik})\, x_k = 0. \end{aligned}\right\} \qquad (П.3.42)$$

Мы видим, что составляющие собственного вектора x_k определяются из системы линейных однородных уравнений (П.3.42), которая имеет решения, отличные

от нулевых, только тогда, когда определитель системы равен нулю, т. е.

$$\begin{vmatrix} \alpha_{11}-\lambda, & \alpha_{12}, & \ldots, & \alpha_{1n} \\ \alpha_{21}, & \alpha_{22}-\lambda, & \ldots, & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1}, & \alpha_{n2}, & \ldots, & \alpha_{nn}-\lambda \end{vmatrix} = 0, \qquad (\text{П.3.43})$$

или

$$|\alpha_{ik}-\lambda\delta_{ik}| = 0. \qquad (\text{П.3.43а})$$

Уравнение (П.3.43) n-й степени относительно λ называется *характеристическим* или *секулярным*; n корней этого уравнения $\lambda_1, \lambda_2, \ldots, \lambda_n$ (из которых некоторые могут совпадать) определяют n собственных значений матрицы $\boldsymbol{\alpha}$. Каждому собственному значению λ_i соответствует собственный вектор $\boldsymbol{x}^{(i)}$ (или несколько собственных векторов; в последнем случае собственное значение называется *вырожденным*). Докажем важную теорему о собственных значениях и собственных векторах эрмитовой матрицы $\boldsymbol{H}$. Пусть

$$\boldsymbol{H}\boldsymbol{x}^{(1)} = \lambda_1 \boldsymbol{x}^{(1)}, \qquad (\text{П.3.44})$$

$$\boldsymbol{H}\boldsymbol{x}^{(2)} = \lambda_2 \boldsymbol{x}^{(2)}, \qquad (\text{П.3.44а})$$

где $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$ и λ_1, λ_2 — два собственных вектора и два собственных значения матрицы $\boldsymbol{H}$.

В проекциях на координатные оси имеем

$$\sum_k H_{ik} x_k^{(1)} = \lambda_1 x_i^{(1)}, \qquad (\text{П.3.45})$$

$$\sum_k H_{ik} x_k^{(2)} = \lambda_2 x_i^{(2)}. \qquad (\text{П.3.45а})$$

Умножим обе части (П.3.45) на $x_i^{(2)*}$ и просуммируем по i, тогда

$$\sum_{k,\,i} H_{ik} x_k^{(1)} x_i^{(2)*} = \lambda_1 \sum_i x_i^{(1)} x_i^{(2)*}.$$

Возьмем уравнение комплексно-сопряженное с (П.3.45а), умножим его на $x_i^{(1)}$ и просуммируем по i

$$\sum_{k\ i} H_{ik}^* x_k^{(2)*} x_i^{(1)} = \lambda_2^* \sum_i x_i^{(2)*} x_i^{(1)}.$$

Так как $\boldsymbol{H}$ — эрмитова матрица, то $H_{ik}^* = H_{ki}$, но тогда левые части двух последних уравнений равны (во втором можно поменять немые индексы i и k) и, следовательно,

$$(\lambda_1 - \lambda_2^*) \sum_i x_i^{(2)*} x_i^{(1)} = 0. \qquad (\text{П.3.46})$$

Предположим сначала, что $\lambda_1 = \lambda_2$, а $\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(2)} \neq 0$; тогда из (П.3.46) следует

$$\lambda_1 = \lambda_2^* = \lambda_1^* \qquad (\text{П.3.47})$$

(последнее равенство следует из того, что $\lambda_1 = \lambda_2$), т. е. *собственные значения* эрмитовых матриц вещественны. Предположим, напротив, что $\lambda_1 \neq \lambda_2 = \lambda_2^*$; тогда из (П.3.46) следует

$$\sum_i x_i^{(2)*} x_i^{(1)} = (\boldsymbol{x}^{(2)}, \boldsymbol{x}^{(1)}) = 0, \qquad (\text{П.3.48})$$

т. е. собственные векторы эрмитовых матриц (относящиеся к разным собственным значениям) взаимно ортогональны. Если собственное значение λ *вырож-*

дено и ему принадлежат, например, три собственных вектора $x^{(1)}$, $x^{(2)}$ и $x^{(3)}$, то собственные векторы, соответствующие λ, тоже могут быть сделаны взаимно ортогональными посредством *процедуры ортогонализации Грама—Шмидта.* Ясно, что любая линейная комбинация векторов $x^{(1)}$, $x^{(2)}$ и $x^{(3)}$ — тоже собственный вектор, принадлежащий λ. Положим $y^{(1)}=x^{(1)}$, $y^{(2)}=x^{(2)}+p_1y^{(1)}$. Выберем p_1 так, чтобы $y^{(1)}$ было ортогонально $y^{(2)}$, т. е.

$$(y^{(1)}, y^{(2)})=0=(y^{(1)}, x^{(2)})+p_1(y^{(1)}, y^{(1)}),$$

откуда

$$p_1=-\frac{(y^{(1)}, x^{(2)})}{(y^{(1)}, y^{(1)})}.$$

Теперь положим $y^{(3)}=x^{(3)}+p_2y^{(1)}+p_3y^{(2)}$ и выберем p_2 и p_3 так, чтобы $y^{(3)}$ было ортогонально $y_1^{(1)}$ и $y^{(2)}$. В результате мы сконструируем три взаимно ортогональных собственных вектора $y^{(i)}$ ($i=1, 2, 3$), относящихся к λ. Конечно, эту процедуру можно распространить на произвольную кратность вырождения собственного значения λ.

6. Введем важное обобщение о произведении матриц разного ранга.

Пусть квадратная матрица $\boldsymbol{\alpha}$ n-го ранга, а квадратная матрица $\boldsymbol{\beta}$ m-го ранга ($n\neq m$). *Прямым произведением* матриц $\boldsymbol{\alpha}$ и $\boldsymbol{\beta}$, обозначаемым через $\boldsymbol{\alpha}\times\boldsymbol{\beta}$, называется матрица с элементами

$$(\boldsymbol{\alpha}\times\boldsymbol{\beta})_{ij,\,kl}=\alpha_{ik}\beta_{jl}. \qquad (\text{П.3.49})$$

Здесь двойной индекс ij, в котором i пробегает значения от 1 до n, j пробегает значения от 1 до m, нумерует строки матрицы $\boldsymbol{\alpha}\times\boldsymbol{\beta}$ в соответствии с некоторым, произвольно выбранным правилом. Например, 1-й, 2-й, 3-й, ..., (nm)-й строке могут соответствовать значения индекса ij: 11, 12, 13, ..., 1m, 21, 22, 23, ..., 2m, ..., n1, n2, ..., nm. Нумерация столбцов kl производится обязательно по тому же правилу (Е. Вигнер говорит, что *разметка* строк и столбцов должна быть одинаковой). Всякое изменение разметки сводится к перестановке строк с одновременной перестановкой столбцов. Можно показать, что это является несущественным для доказываемых ниже свойств.

Теорема: Пусть имеются две матрицы $\boldsymbol{\alpha}$ и $\boldsymbol{\alpha}'$ n-го ранга и две матрицы $\boldsymbol{\beta}$ и $\boldsymbol{\beta}'$ m-го ранга, тогда

$$(\boldsymbol{\alpha}\times\boldsymbol{\beta})(\boldsymbol{\alpha}'\times\boldsymbol{\beta}')=\boldsymbol{\alpha}\boldsymbol{\alpha}'\times\boldsymbol{\beta}\boldsymbol{\beta}'. \qquad (\text{П.3.50})$$

Здесь в левой части равенства мы вначале вычисляем прямые произведения $\boldsymbol{\alpha}\times\boldsymbol{\beta}$ и $\boldsymbol{\alpha}'\times\boldsymbol{\beta}'$, а затем обычное произведение полученных таким образом матриц $n\cdot m$-го ранга; в правой части мы вначале вычисляем обычные произведения матриц $\boldsymbol{\alpha}\boldsymbol{\alpha}'$ и $\boldsymbol{\beta}\boldsymbol{\beta}'$, а затем их прямое произведение.

Доказательство. Элемент с индексом (ij, kl) левой части равен

$$\{(\boldsymbol{\alpha}\times\boldsymbol{\beta})(\boldsymbol{\alpha}'\times\boldsymbol{\beta}')\}_{ij,\,kl}=\sum_{r\,s}(\boldsymbol{\alpha}\times\boldsymbol{\beta})_{ij,\,rs}(\boldsymbol{\alpha}'\times\boldsymbol{\beta}')_{rs,\,kl}=\sum_{r\,s}\alpha_{ir}\beta_{js}\alpha'_{rk}\beta'_{sl}, \qquad (\text{П.3.51})$$

где первое равенство выражает закон обычного матричного умножения (П.3.15) (т. е. суммирование по «внутреннему» индексу rs), а второе равенство основано на (П.3.49).

Тот же элемент с индексом (ij, kl) правой части (П.3.50) равен

$$\{\boldsymbol{\alpha}\boldsymbol{\alpha}'\times\boldsymbol{\beta}\boldsymbol{\beta}'\}_{ij,\,kl}=(\boldsymbol{\alpha}\boldsymbol{\alpha}')_{ik}(\boldsymbol{\beta}\boldsymbol{\beta}')_{jl}=\sum_{r\,s}\alpha_{ir}\alpha'_{rk}\beta_{js}\beta'_{sl},$$

что совпадает с (П.3.51). Таким образом, теорема (П.3.50) доказана.

Вычислим еще шпур матрицы, равной прямому произведению матриц

$$\mathrm{Sp}\,(\boldsymbol{\alpha}\times\boldsymbol{\beta})=\sum_{ij}(\boldsymbol{\alpha}\times\boldsymbol{\beta})_{ij,\,ij}=\sum_{ij}\alpha_{ii}\beta_{jj}=\sum_i\alpha_{ii}\sum_j\beta_{jj}=\mathrm{Sp}\,\boldsymbol{\alpha}\cdot\mathrm{Sp}\,\boldsymbol{\beta}, \qquad (\text{П.3.52})$$

т. е. шпур прямого произведения двух матриц равен произведению шпуров самих матриц.

Приложение 4

Рассмотрим свойства идеального ферми-газа при абсолютном нуле температуры (Я. И. Френкель). Пусть в объеме ΔV содержится ΔN электронов в наинизшем энергетическом состоянии. Принцип Паули требует, чтобы элементарная ячейка фазового пространства $(\Delta x\,\Delta y\,\Delta z)\,\Delta p_x\,\Delta p_y\,\Delta p_z=(2\pi\hbar)^3$ содержала не больше двух электронов (с противоположно направленными спинами).

В наинизшем энергетическом состоянии ΔN электронов заполнят в пространстве импульсов шар радиуса p_0, который определится из условия

$$2\frac{\text{полный фазовый объем}}{(2\pi\hbar)^3}=2\frac{\Delta V\dfrac{4\pi}{3}p_0^3}{(2\pi\hbar)^3}=\Delta N, \tag{П.4.1}$$

откуда

$$p_0=2\pi\hbar\left(\frac{3n}{8\pi}\right)^{1/3}, \tag{П.4.2}$$

где $n=\Delta N/\Delta V$ — концентрация электронов. Максимальная кинетическая энергия электронов в точке с концентрацией n равна

$$\varepsilon_0=\frac{1}{2m}p_0^2=\frac{(2\pi\hbar)^2}{2m}\left(\frac{3n}{8\pi}\right)^{2/3}. \tag{П.4.3}$$

Число квантовых состояний (статистический вес) в 1 $см^3$ в интервале импульса между p и $p+dp$ или кинетической энергии между ε и $\varepsilon+d\varepsilon$ равно

$$\begin{aligned} d_pn&=2\frac{1\ см^3\times\text{объем шаров слоя }dp}{(2\pi\hbar)^3}=\frac{8\pi p^2\,dp}{(2\pi\hbar)^3};\\ d_\varepsilon n&=\frac{8\sqrt{2}\pi m^{3/2}}{(2\pi\hbar)^3}\sqrt{\varepsilon}\,d\varepsilon. \end{aligned} \tag{П.4.4}$$

Здесь $\varepsilon=p^2/2m$. На рис. П.2 изображены плотности распределения электронов (квантовых состояний) на единичные интервалы импульса и энергии.

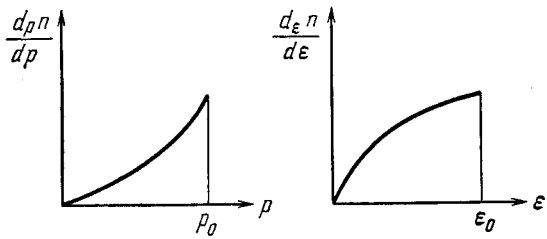


Рис. П. 2.

Вычислим плотность кинетической энергии электронов (т. е. кинетическую энергию в расчете на 1 $см^3$):

$$\mathscr{E}_k=\int\frac{1}{2m}p^2d_pn=\frac{8\pi}{2m(2\pi\hbar)^3}\int_0^{p_0}p^4\,dp=\frac{3^{5/3}\pi^{4/3}\hbar^2}{10m}n^{5/3}. \tag{П.4.5}$$

Приложение 5

Понятие о нормальных (главных) координатах механической системы. Рассмотрим простую механическую систему с двумя степенями свободы, на которой может быть наглядно проиллюстрировано понятие о *нормальных* или *главных координатах*. Представим себе две частицы с одинаковыми (для простоты) массами m, которые могут двигаться вдоль оси x. На 1-ю и 2-ю частицы действуют квазиупругие силы $-\beta u_1$ и $-\beta u_2$, притягивающие их к положениям равновесия O_1 и O_2 (u_1 и u_2—отклонение частиц от центров O_1 и O_2). Кроме того, частицы взаимодействуют друг с другом с квазиупругой силой, равной $\pm\gamma(u_1-u_2)$. Уравнения движения имеют вид

$$\begin{aligned} m\ddot{u}_1 &= -\beta u_1 - \gamma(u_1-u_2), \\ m\ddot{u}_2 &= -\beta u_2 - \gamma(u_2-u_1). \end{aligned} \tag{П.5.1}$$

Ищем решение этой совместной системы дифференциальных уравнений в комплексной форме

$$u_1 = A_1 e^{i\omega t}, \quad u_2 = A_2 e^{i\omega t}. \tag{П.5.2}$$

Подставляя (П.5.2) в (П.5.1) и сокращая обе части уравнений на множитель $e^{i\omega t}$, получим линейную однородную алгебраическую систему уравнений для амплитуд A_1 и A_2

$$\begin{aligned} \left(\omega_A^2-\omega^2\right)A_1 - \omega_B^2 A_2 &= 0, \\ -\omega_B^2 A_1 + \left(\omega_A^2-\omega^2\right)A_2 &= 0, \end{aligned} \tag{П.5.3}$$

где $m\omega_A^2=\beta+\gamma$ и $m\omega_B^2=\gamma$.

Как мы видели в аналогичном случае § 3, п. 2, система (П.5.3) определяет только отношение A_1/A_2. Из (П.5.3)

$$\frac{A_1}{A_2} = \frac{\omega_B^2}{\omega_A^2-\omega^2} = \frac{\omega_A^2-\omega^2}{\omega_B^2}. \tag{П.5.4}$$

Решая это *квадратное характеристическое уравнение* для ω^2, получим два корня

$$\omega_1^2 = \omega_A^2+\omega_B^2 = \frac{\beta+2\gamma}{m}, \quad \omega_2^2 = \omega_A^2-\omega_B^2 = \frac{\beta}{m}. \tag{П.5.5}$$

Из (П.5.4) следует, что каждый из этих корней дает для отношения A_1/A_2

$$\left(\frac{A_1}{A_2}\right)_{\omega=\omega_1} = -1, \quad \text{т. е. } A_1^{(1)} = -A_2^{(1)} \equiv c_1,$$

$$\left(\frac{A_1}{A_2}\right)_{\omega=\omega_2} = +1, \quad \text{т. е. } A_1^{(2)} = +A_2^{(2)} \equiv c_2.$$

Отсюда и из линейности уравнений (П.5.1) следует, что общее решение имеет вид

$$u_1 = c_1 e^{i\omega_1 t} + c_2 e^{i\omega_2 t}, \quad u_2 = -c_1 e^{i\omega_1 t} + c_2 e^{i\omega_2 t}. \tag{П.5.6}$$

Мы видим, что координаты u_1 и u_2 состоят из сумм, каждое слагаемое которых меняется по времени по гармоническому закону со своей частотой. Возникает вопрос, нельзя ли выбрать в системе такие координаты (*нормальные*), чтобы каждая из них зависела от времени гармонически. В данном случае это сделать очень просто. Из (П.5.6) непосредственно видно, что

$$(u_1+u_2) \sim e^{i\omega_2 t}; \quad (u_1-u_2) \sim e^{i\omega_1 t};$$

определим нормальные координаты, полагая

$$q_1=\frac{u_1-u_2}{\sqrt{2}}, \quad q_2=\frac{u_1+u_2}{\sqrt{2}} \tag{П.5.7}$$

(постоянный множитель $\sqrt{2}$ выбран для удобства). Решая (П.5.7) относительно u_1 и u_2, получим

$$u_1=\frac{\sqrt{2}}{2}(q_1+q_2), \quad u_2=\frac{\sqrt{2}}{2}(q_2-q_1). \tag{П.5.8}$$

Выразим теперь кинетическую энергию $\mathcal{T}$ и потенциальную Ф через нормальные координаты q_1 и q_2 и обобщенные скорости $\dot{q}_1$ и $\dot{q}_2$. Кинетическая энергия

$$\mathcal{T}=\frac{m}{2}(\dot{u}_1^2+\dot{u}_2^2)=\frac{m}{2}(\dot{q}_1^2+\dot{q}_2^2). \tag{П.5.9}$$

Потенциальная энергия

$$\Phi=\frac{1}{2}\{\beta(u_1^2+u_2^2)+\gamma(u_1-u_2)^2\}=\frac{1}{2}\{(\beta+\gamma)(u_1^2+u_2^2)-2\gamma u_1u_2\}.$$

В этом легко убедиться, вычисляя силы $-\partial\Phi/\partial u_1$ и $-\partial\Phi/\partial u_2$, действующие на частицы. Подставляя в Ф нормальные координаты по формулам (П.5.8), получим

$$\Phi=\frac{m}{2}(\omega_1^2q_1^2+\omega_2^2q_2^2). \tag{П.5.10}$$

Таким образом, функция Лагранжа

$$\mathcal{L}=\mathcal{T}-\Phi=\frac{m}{2}(\dot{q}_1^2+\dot{q}_2^2)-\frac{m}{2}(\omega_1^2q_1^2+\omega_2^2q_2^2). \tag{П.5.11}$$

Мы видим, что $\mathcal{L}$ в нормальных координатах q_i и соответствующих обобщенных скоростях $\dot{q}_i$ свелась к сумме квадратов (не содержит смешанных членов, например, пропорциональных q_1q_2). Легко видеть, что это условие является необходимым и достаточным для того, чтобы каждая из координат гармонически зависела от времени. Напишем уравнения Лагранжа второго рода $\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}-\frac{\partial\mathcal{L}}{\partial q_i}=0$ для функции Лагранжа (П.5.11). Имеем

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i}=m_i\ddot{q}_i \quad \text{и} \quad \frac{\partial\mathcal{L}}{\partial q_i}=-m\omega_i^2q_i,$$

откуда

$$\ddot{q}_i+\omega_i^2q_i=0 \quad (i=1,\ 2). \tag{П.5.12}$$

Мы видим, что $q_i\sim e^{i\omega_i t}$ в согласии с тем, что мы имели выше. В силу того, что Ф, а следовательно, и функция Лагранжа $\mathcal{L}=\mathcal{T}-\Phi$ содержат в переменных u_i смешанный член $-2\gamma u_1u_2$ в уравнениях движения (П.5.1) переменные u_1 и u_2 не разделяются. Поэтому каждая из координат u_i зависит от времени не гармонически, а более сложно.

Введем обобщенные импульсы, соответствующие нормальным координатам q_i, $p_i=\partial\mathcal{L}/\partial\dot{q}_i=m\dot{q}_i$. Функция Гамильтона нашей системы

$$\mathcal{H}(q,\ p)=\mathcal{T}+\Phi=\frac{1}{2m}(p_1^2+p_2^2)+\frac{1}{2m}(\omega_1^2q_1^2+\omega_2^2q_2^2), \tag{П.5.13}$$

т. е. тоже содержит только квадраты канонических переменных q и p. Можно сказать, что задача определения нормальных координат системы состоит в приведении функции Гамильтона $\mathcal{H}$ к сумме квадратов q и p.

Приложение 6

Вычислим две суммы

$$L=\sum_{a_n} e^{iqa_n} \quad \text{и} \quad M=\sum_{q} e^{iqa_n}, \tag{П.6.1}$$

где вектор прямой решетки

$$a_n=n_1a_1+n_2a_2+n_3a_3=\sum_{k=1}^{3} n_ka_k \quad (n_k=1,\ 2,\ 3,\ \ldots,\ G) \tag{П.6.1а}$$

и волновой вектор

$$q=\frac{1}{G}(g_1b_1+g_2b_2+g_3b_3)=\sum_{i=1}^{3}\frac{g_i}{G}\,b_i\left(-\frac{G}{2}<g_i<\frac{G}{2}\right)^{1)}, \tag{П.6.1б}$$

как это следует из условий цикличности Борна—Кармана (§ 5, п. 6). Сумма L является трехмерным обобщением выражения (4.5). Первая сумма

$$L=\sum_{n_1n_2n_3}\exp\left[i\frac{1}{G}\sum_{i=1}^{3}g_ib_i\sum_{k=1}^{3}n_ka_k\right]=\sum_{n_1n_2n_3}\exp\left[\frac{2\pi i}{G}\sum_{i=1}^{3}g_in_i\right], \tag{П.6.2}$$

где мы использовали условие: $b_ia_k=2\pi\delta_{ik}$ (см. (I.3.9)).

1) $q\neq 0$, тогда хотя бы одно $g_i\neq 0$ и, следовательно, соответствующее

$$l_i\equiv\exp\left(\frac{2\pi i}{G}g_i\right)\neq 1. \tag{П.6.3}$$

В этом случае

$$L=\sum_{n_1n_2n_3} l_1^{n_1}l_2^{n_2}l_3^{n_3}\approx\sum_{n_i} l_i^{n_i},$$

но

$$\sum_{n_i=1}^{G} l_i^{n_i}=l_i+l_i^2+\ldots+l_i^G=\frac{l_i(1-l_i^G)}{1-l_i}=0,$$

так как $l_i^G=\exp 2\pi i g_i=1$, а знаменатель, согласно (П.6.3), отличен от нуля. Таким образом, в этом случае $L=0$.

2) $q=0$, т. е. все $g_i=0$; тогда все $l_i=1$ и $L=\sum_{n_1n_2n_3}1^{n_1}1^{n_2}1^{n_3}=\sum_{n_1n_2n_3}1=$ $=G^3=N=$числу ячеек основной области кристалла.

Ясно, что случай $q=b_g$— вектору обратной решетки эквивалентен случаю $q=0$.

Таким образом,

$$L=\sum_{a_n} e^{iqa_n}=N\delta_{qb_g}, \tag{П.6.4}$$

где, в частности, $b_g=0$.

Вторая сумма

$$M=\sum_{g_1g_2g_3}\exp\left[\frac{2\pi i}{G}\sum_{i=1}^{3}g_in_i\right]. \tag{П.6.5}$$

---

1) Удобно считать, что G—большое нечетное число.

Обозначим через

$$m_i \equiv \exp\left(\frac{2\pi i}{G} n_i\right), \tag{П.6.6}$$

тогда

$$M = \sum_{g_1 g_2 g_3} m_1^{g_1} m_2^{g_2} m_3^{g_3}. \tag{П.6.7}$$

Рассмотрим опять два случая:

1) $a_n \neq 0$, тогда хотя бы одно $n_i \neq 0$ соответствует $m_i \neq 1$.

В этом случае сумма M имеет множитель

$$\sum_{g_i=-\frac{G-1}{2}}^{\frac{G-1}{2}} m_i^{g_i} = m_i^{-\frac{G-1}{2}} + m_i^{-\frac{G-3}{2}} + \ldots + m_i^{\frac{G-1}{2}} =$$

$$= m_i^{-\frac{G-1}{2}} \left[1 + m_i + \ldots + m_i^{G-1}\right] = m_i^{-\frac{G-1}{2}} \frac{1 - m_i^G}{1 - m_i} = 0,$$

так числитель $1 - m_i^G = 1 - \exp(2\pi i n_i) = 1 - 1 = 0$, а знаменатель отличен от нуля.

2) $a_n = 0$ тогда все $n_i = 0$ и, следовательно, все $m_i = 1$, в этом случае, так же как и в предыдущем, сумма $M = N$. Таким образом,

$$M = \sum_q e^{iqa_n} = N\delta_{a_n 0}. \tag{П.6.8}$$

Приложение 7

1. Приведем некоторые определенные интегралы, часто встречающиеся при изучении теории полупроводников [1]).

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = 2\int_0^{\infty} e^{-\alpha x^2} dx = \frac{\sqrt{\pi}}{\alpha^{1/2}}, \tag{П.7.1}$$

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} x^2 dx = 2\int_0^{\infty} e^{-\alpha x^2} x^2 dx = \frac{1}{2}\frac{\sqrt{\pi}}{\alpha^{3/2}}, \tag{П.7.2}$$

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} x^4 dx = 2\int_0^{\infty} e^{-\alpha x^2} x^4 dx = \frac{3}{4}\frac{\sqrt{\pi}}{\alpha^{5/2}}, \tag{П.7.3}$$

$$\int_0^{\infty} e^{-\alpha x^2} x\, dx = \frac{1}{2\alpha}, \tag{П.7.4}$$

$$\int_0^{\infty} e^{-\alpha x^2} x^3 dx = \frac{1}{2\alpha^2}, \tag{П.7.5}$$

$$\int_0^{\infty} e^{-\alpha x^2} x^5 dx = \frac{1}{\alpha^3}. \tag{П.7.6}$$

[1]) Вывод этих формул можно найти, например, у Смирнова В. И., т. 2, с. 258.

2. Определим гамма-функцию $\Gamma(z)$ посредством равенства

$$\Gamma(z)=\int_0^\infty x^{z-1}e^{-x}\,dx. \qquad (\text{П.7.7})$$

Определенный интеграл в правой части, взятый по вещественной переменной x, зависит от z, как от параметра. Равенство (П.7.7) определяет гамма-функцию для любых комплексных значений z с положительной вещественной частью.

Выведем важное рекуррентное соотношение

$$\Gamma(z+1)=z\Gamma(z). \qquad (\text{П.7.8})$$

Интегрируя по частям, получим

$$\Gamma(z+1)=\int_0^\infty x^z e^{-x}\,dx=-\int_0^\infty x^z\,de^{-x}=-x^z e^{-x}\Big|_0^\infty+\int_0^\infty zx^{z-1}e^{-x}\,dx=$$

$$=z\int_0^\infty x^{z-1}e^{-x}\,dx=z\Gamma(z).$$

Для того, чтобы определить значение $\Gamma(z)$ для целых и полуцелых значений z, вычислим

$$\Gamma(1)=\int_0^\infty e^{-x}\,dx=1 \qquad (\text{П.7.9})$$

и

$$\Gamma\left(\frac{1}{2}\right)=\int_0^\infty x^{1/2}e^{-x}\,dx=\sqrt{\pi}, \qquad (\text{П.7.10})$$

где мы ввели переменную интегрирования $t=x^{1/2}$.

Используя рекуррентное соотношение (П.7.8) и значения $\Gamma(1)$ и $\Gamma(1/2)$, получим

$$\Gamma(2)=1,\ \Gamma(3)=1\cdot 2,\ \Gamma(4)=1\cdot 2\ 3,\ \ldots,\ \Gamma(n)=(n-1)!, \qquad (\text{П.7.11})$$

$$\Gamma\left(\frac{3}{2}\right)=\frac{1}{2}\sqrt{\pi},\quad \Gamma\left(\frac{5}{2}\right)=\frac{3}{2}\frac{1}{2}\sqrt{\pi},\quad \Gamma\left(\frac{7}{2}\right)=\frac{5}{2}\frac{3}{2}\frac{1}{2}\sqrt{\pi},$$

$$\Gamma\left(\frac{2n+1}{2}\right)=\frac{(2n-1)!!}{2^n}\sqrt{\pi}, \qquad (\text{П.7.12})$$

где $(2n-1)!!$ — произведение последовательных нечетных чисел от 1 до $(2n-1)$.

Приложение 8

1. Для упрощения вычислений выведем формулу (2.9), т. е. вычислим энергию $\mathscr{E}$ для системы, состоящей из $N=2$ электронов. Обобщение на произвольное число N электронов не представляет принципиальных трудностей, для этого необходимо только использовать некоторые хорошо известные свойства определителей.

Используя волновую функцию (2.8) и гамильтониан (2.1а) для системы двух электронов, получим

$$\mathscr{E}=\int \Phi^*\hat{\mathscr{H}}\Phi\,d\tau_1\,d\tau_2=\frac{1}{2}\int[\varphi_1^*(1)\,\varphi_2^*(2)-\varphi_1^*(2)\,\varphi_2^*(1)]\times$$

$$\times\left\{\hat{\mathscr{H}}_1+\hat{\mathscr{H}}_2+\frac{e^2}{r_{12}}\right\}[\varphi_1(1)\,\varphi_2(2)-\varphi_1(2)\,\varphi_2(1)]\,d\tau_1\,d\tau_2, \qquad (\text{П.8.1})$$

где $\hat{\mathscr{H}}_i$ $(i=1,2)$ определяются выражением (2.1б).

Раскрывая скобки под интегралом, получим

$$\mathscr{E}=\frac{1}{2}\left\{\int\varphi_1^*(1)\,\hat{\mathscr{H}}_1\varphi_1(1)\,d\tau_1+\int\varphi_1^*(2)\,\hat{\mathscr{H}}_2\varphi_1(2)\,d\tau_2+\right.$$
$$+\int\varphi_2^*(1)\,\hat{\mathscr{H}}_1\varphi_2(1)\,d\tau_1+\int\varphi_2^*(2)\,\hat{\mathscr{H}}_2\varphi_2(2)\,d\tau_2+$$
$$+\int|\varphi_1(1)|^2\frac{e^2}{r_{12}}|\varphi_2(2)|^2\,d\tau_1\,d\tau_2+\int|\varphi_1(2)|^2\frac{e^2}{r_{12}}|\varphi_2(1)|^2\,d\tau_1\,d\tau_2-$$
$$-\int\varphi_1^*(1)\,\varphi_2(1)\frac{e^2}{r_{12}}\varphi_1(2)\,\varphi_2^*(2)\,d\tau_1\,d\tau_2-$$
$$\left.-\int\varphi_1^*(2)\,\varphi_2(2)\frac{e^2}{r_{12}}\varphi_1(1)\,\varphi_2^*(1)\,d\tau_1\,d\tau_2\right\}.$$

Здесь 1-й интеграл равен 2-му, 3-й—4-му, 5-й—6-му и 7-й—8-му, так как как эти пары интегралов различаются переменными интегрирования, а величина определенного интеграла от переменной интегрирования не зависит

$$\mathscr{E}=\sum_i^{1,2}\int\varphi_i^*(1)\,\hat{\mathscr{H}}_1\varphi_i(1)\,d\tau_1+\frac{1}{2}\sum_{i\neq j}^{1,2}\int|\varphi_i(1)|^2\frac{e^2}{r_{12}}|\varphi_j(2)|^2\,d\tau_1\,d\tau_2-$$
$$-\frac{1}{2}\sum_{i\neq j}^{1,2}\int\varphi_i^*(1)\,\varphi_j(1)\frac{e^2}{r_{12}}\varphi_i(2)\,\varphi_j^*(2)\,d\tau_1\,d\tau_2, \qquad (\text{П.8.2})$$

что совпадает с (2.9) в случае $N=2$.

2. Из (2.9а) следует

$$\delta\mathscr{E}=\sum_i^{1,N}\int d\boldsymbol{r}_1\,\delta\psi_{n_i}^*(\boldsymbol{r}_1)\left\{\hat{\mathscr{H}}_1\psi_{n_i}(\boldsymbol{r}_1)+\left[\sum_j{}'\int\frac{e^2|\psi_{n_j}(\boldsymbol{r}_2)|^2\,d\boldsymbol{r}_2}{r_{12}}\psi_{n_i}(\boldsymbol{r}_1)-\right.\right.$$
$$\left.\left.-\sum_{\substack{\text{(по параллельным}\\ \text{спинам)}}}{}'\left(\int\frac{e^2\psi_{n_i}(\boldsymbol{r}_2)\,\psi_{n_j}^*(\boldsymbol{r}_2)\,d\boldsymbol{r}_2}{r_{12}}\right)\psi_{n_j}(\boldsymbol{r}_1)\right]\right\}. \qquad (\text{П.8.3})$$

Множитель $^1/_2$ перед двойными суммами в (2.9а) исчезает, так как при варьировании функций $\psi_{n_i}^*$ функция с данными n_i встречается дважды: один раз при суммировании по i другой раз—по j. Дополнительное условие (2.10а) учтем методом неопределенных множителей Лагранжа. Для этого умножим (2.10а) на $-\lambda_{ij}$, просуммируем по i и j, прибавим к (П.8.3) и приравняем нулю

$$\sum_i^{1,N}\int d\boldsymbol{r}_1\,\delta\psi_{n_i}^*(\boldsymbol{r}_1)\left\{\hat{\mathscr{H}}_1\psi_{n_i}(\boldsymbol{r}_1)+\ldots-\sum_j\lambda_{ij}\psi_{n_j}(\boldsymbol{r}_1)\right\}=0. \qquad (\text{П.8.4})$$

Ввиду произвольности вариаций $\delta\psi_{n_i}^*$ фигурная скобка в (П.8.4) для всех значений i равна нулю. Можно показать, что решения всегда могут быть выбраны так, что матрица λ_{ij} диагональна. Обозначая $\lambda_{ii}=\mathscr{E}_{n_i}$, получим уравнение (2.11).

Приложение 9

Подставим в выражение (2.11б) для $\mathcal{U}_{\text{eff}}$ блоховские функции (3.1), считая $n_i \equiv k$ и $n_j = k'$

$$\mathcal{U}_{\text{eff}}(r_1) = \sum_{k'}{}' \int \frac{e^2 |u_{k'}(r_2)|^2}{|r_2 - r_1|} dr_2 - \sum_{k'}{}' \frac{u_{k'}(r_1) e^{ik'r_1}}{u_k(r_1) e^{ikr_1}} \int \frac{e^2 u_k(r_2) e^{ikr_2} u^*_{k'}(r_2) e^{-ik'r_2}}{|r_2 - r_1|} dr_2.$$

Заменим в этом выражении r_1 на $r_1 + a_n$ и r_2 на $r_2 + a_n$; последнее сводится к замене переменной интегрирования. Используя свойства периодичности (3.1а) для функций u_k и $u_{k'}$ и сокращая во второй сумме справа множитель $e^{i(k'-k)a_n}$, получим, что $\mathcal{U}_{\text{eff}}(r_1 + a_n)$ равно исходному выражению, определяющему $\mathcal{U}_{\text{eff}}(r_1)$, т. е. докажем периодичность $\mathcal{U}_{\text{eff}}$.

Приложение 10

При подстановке (3.1) в уравнение Шредингера (3.7) необходимо вычислить оператор Лапласа $\nabla^2 \equiv \operatorname{div}\operatorname{grad} \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ от произведения двух функций $u_k(r) e^{ikr}$, зависящих от $(r \equiv x, y, z)$. Легко убедиться в справедливости следующей формулы:

$$\nabla^2 \{f(r) \varphi(r)\} = \operatorname{div}\operatorname{grad}(f\varphi) = f\nabla^2\varphi + \varphi\nabla^2 f + 2(\operatorname{grad} f, \operatorname{grad}\varphi). \quad \text{(п.10.1)}$$

Полагая $f = u_k(r)$ и $\varphi = e^{ikr}$, имеем

$$\operatorname{grad} f = \operatorname{grad} u_k, \quad \operatorname{grad}\varphi = ike^{ikr},$$
$$\nabla^2 f = \nabla^2 u_k, \qquad \nabla^2\varphi = -k^2 e^{ikr}.$$

Подставляя эти значения в формулу (П.10.1), получим

$$\nabla^2(f\varphi) = \nabla^2(u_k e^{ikr}) = \{\nabla^2 u_k + 2i(k \operatorname{grad} u_k) - k^2 u_k\} e^{ikr}. \quad \text{(П.10.2)}$$

Используя это выражение, легко получить (3.8).

Приложение 11

1. Рассмотрим простейшие свойства, связанные с понятием тензора. Вектор обычно определяют как величину, которая в отличие от скаляра характеризуется не только своим численным значением, но и направлением в пространстве. Простейшими примерами скаляра и вектора являются, например, масса частицы m и радиус-вектор r, определяющий ее положение в пространстве. Как мы сейчас покажем, определение вектора нуждается в уточнении, которое позволит нам одновременно дать определение тензора.

Рассмотрим две прямоугольные координатные системы (x_1, x_2, x_3) и (x'_1, x'_2, x'_3), имеющие общее начало координат O. Косинусы углов между осями обеих систем обозначим через $\alpha_{ik} = \cos(\widehat{x_i x'_k})$, где i и k независимо пробегают значения 1, 2 и 3.

Поставим вопрос, является ли проекция радиус-вектора r на некоторую ось x_i скаляром. С одной стороны, проекция характеризуется только некото-

рым численным значением, с другой — при переходе к штрихованной координатной системе — $x_i' \neq x_i$, в то время как скалярная масса m одинакова в обеих системах.

Из аналитической геометрии известно[1]), что проекции радиуса-вектора $\boldsymbol{r}$ в обеих координатных системах связаны соотношениями

$$\begin{aligned} x_1' &= x_1\alpha_{11} + x_2\alpha_{12} + x_3\alpha_{13}, \\ x_2' &= x_1\alpha_{21} + x_2\alpha_{22} + x_2\alpha_{23}, \\ x_3' &= x_1\alpha_{31} + x_2\alpha_{32} + x_3\alpha_{33} \end{aligned} \qquad (\text{П.11.1})$$

или короче

$$x_i' = \sum_{k=1}^{3} \alpha_{ik} x_k \quad (i = 1,\ 2,\ 3). \qquad (\text{П.11.2а})$$

Аналогично

$$x_i = \sum_{k=1}^{3} \alpha_{ki} x_k' \quad (i = 1,\ 2,\ 3). \qquad (\text{П.11.2б})$$

Дадим теперь строгое определение понятия вектора. Будем под вектором понимать совокупность трех величин A_i $(i = 1, 2, 3)$, которые при переходе от одной координатной системы к другой преобразуются по законам (П.11.2а) и (П.11.2б), т. е.

$$A_i' = \sum_k \alpha_{ik} A_k, \quad A_i = \sum_k \alpha_{ki} A_k'. \qquad (\text{П.11.2в})$$

2. Рассмотрим, по какому закону преобразуются при переходе от одной координатной системы к другой произведения вида $A_iB_k = T_{ik}$ из составляющих двух векторов $\boldsymbol{A}$ и $\boldsymbol{B}$.

Очевидно,

$$T_{ik}' = A_i'B_k' = \sum_l \alpha_{il} A_l \sum_m \alpha_{km} B_m = \sum_{l,\ m} \alpha_{il}\alpha_{km} A_l B_m,$$

т. е.

$$T_{ik}' = \sum_{l,\ m} \alpha_{il}\alpha_{km} T_{lm}. \qquad (\text{П.11.3})$$

Всякая совокупность девяти величин

$$(T_{ik}) = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}, \qquad (\text{П.11.4})$$

преобразующихся по закону (П.11.3), называется тензором 2-го ранга. С этой точки зрения естественно вектор и скаляр назвать тензорами 1-го и нулевого ранга. Аналогично тензорами 3-го ранга называется совокупность 27 величин T_{ikl}, преобразующихся по закону

$$T_{ikl}' = \sum_{m,\ n,\ p} \alpha_{im}\alpha_{kn}\alpha_{lp} T_{mnp}. \qquad (\text{П.11.5})$$

3. Таким образом, тензор не есть просто совокупность скалярных величин, остающихся постоянными при переходе к новой координатной системе, а есть некоторая совокупность величин, преобразующихся при этом по определенному закону.

В физике тензоры появляются обычно как коэффициенты в соотношениях, связывающих между собой компоненты различных векторов и скаляры с век-

[1]) См. Приложение 3, п. 1.

торами. Рассмотрим примеры тензоров. В изотропной среде дифференциальный закон Ома имеет вид

$$\boldsymbol{j}=\sigma\boldsymbol{E}, \tag{П.11.6}$$

где $\boldsymbol{j}$ — вектор плотности тока, $\boldsymbol{E}$ — вектор напряженности электрического поля и σ — удельная электропроводность вещества в некоторой точке. В проекциях на прямоугольные координатные оси имеем

$$j_x=\sigma E_x,\quad j_y=\sigma E_y,\quad j_z=\sigma E_z. \tag{П.11.6а}$$

В случае анизотропной среды (П.11.6а) обобщается посредством естественного предположения, что каждая составляющая плотности тока — однородная линейная функция всех составляющих напряженности поля, т. е.

$$\begin{aligned} j_x&=\sigma_{xx}E_x+\sigma_{xy}E_y+\sigma_{xz}E_z,\\ j_y&=\sigma_{yx}E_x+\sigma_{yy}E_y+\sigma_{yz}E_z,\\ j_z&=\sigma_{zx}E_x+\sigma_{zy}E_y+\sigma_{zz}E_z \end{aligned} \tag{П.11.7}$$

или короче

$$j_i=\sum_k \sigma_{ik}E_k, \tag{П.11.7а}$$

где индексы i и k принимают значения x, y и z или соответственно 1, 2 и 3.

При переходе к штрихованной координатной системе векторы j_i и E_k преобразуются по закону (П.11.2в). Можно показать, что величины σ_{ik} преобразуются при этом по закону (*П.11.3*), т. е. образуют *тензор 2-го ранга*.

Из (П.11.7а) и (П.11.2в) следует

$$\sum_l \alpha_{li} j'_l=\sum_k \sigma_{ik}\sum_m \sigma_{mk}E'_m.$$

Умножая обе части на α_{ni} и суммируя по i, получим

$$\sum_l\left(\sum_i \alpha_{li}\alpha_{ni}\right)j'_l=\sum_m\left(\sum_{i,\,k}\alpha_{ni}\alpha_{mk}\sigma_{ik}\right)E'_m.$$

Используя для левой части (П.3.1), получим

$$j'_n=\sum_m \sigma'_{nm}E'_m, \tag{П.11.7б}$$

где

$$\sigma'_{nm}=\sum_{i,\,k}\alpha_{ni}\alpha_{mk}\sigma_{ik}. \tag{П.11.7в}$$

Из сравнения (П.11.7в) с (П.11.3) видно, что коэффициенты в (П.11.7) являются компонентами тензора 2-го ранга.

Можно показать, что тензор электропроводности

$$(\sigma_{ik})=\begin{pmatrix}\sigma_{xx} & \sigma_{xy} & \sigma_{xz}\\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz}\\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz}\end{pmatrix} \tag{П.11.8}$$

является симметричным тензором, т. е.

$$\sigma_{ik}=\sigma_{ki}. \tag{П.11.9}$$

Таким образом, компоненты, расположенные симметрично относительно *главной диагонали* тензора σ_{ii}, одинаковы.

В изотропных диэлектриках существует следующая связь между напряженностью электрического поля $\boldsymbol{E}$ и вектором электрической индукции $\boldsymbol{D}$:

$$\boldsymbol{D}=\varepsilon\boldsymbol{E}, \tag{П.11.10}$$

где ε—диэлектрическая постоянная. Для анизотропных сред (кристаллов) это соотношение обобщается следующим образом:

$$D_i = \sum_k \varepsilon_{ik} E_k, \qquad \text{(П.11.10а)}$$

где ε_{ik}—симметричный тензор диэлектрической постоянной.

Энергия электрона в кристалле вблизи ее экстремального значения равна, согласно (3.18),

$$\varepsilon(\boldsymbol{k}) - \varepsilon(\boldsymbol{k}_0) = \Delta\varepsilon = {}^1/_2 \sum_{i,\,l} m_{il}^{-1} p_i p_l, \qquad \text{(П.11.11)}$$

где p_i и p_l—составляющие вектора квазиимпульса электрона, m_{il}^{-1}—величины, определяемые равенством (3.19), i и l—индексы, принимающие значения x, y и z (или 1, 2 и 3). При переходе к новой (штрихованной) координатной системе величина энергии $\Delta\varepsilon$ не меняется, а составляющие вектора $\boldsymbol{p}$ преобразуются по формулам (П.11.2в)

$$p_i = \sum_n \alpha_{ni} p'_n, \quad p_l = \sum_s \alpha_{sl} p'_s.$$

Таким образом,

$$(\Delta\varepsilon) = \frac{1}{2} \sum_{n,\,s} \left(m_{ns}^{-1}\right)' p'_n p'_s, \qquad \text{(П.11.11а)}$$

где

$$\left(m_{ns}^{-1}\right)' = \sum_{i,\,l} \alpha_{ni} \alpha_{sl} m_{il}^{-1}. \qquad \text{(П.11.11б)}$$

Из последнего равенства видно, что величины m_{il}^{-1} при переходе к новой координатной системе преобразуются по закону (П.11.3), т. е. как компоненты тензора 2-го ранга. Из (3.19) видно, что тензор обратной эффективной массы $\left(m_{il}^{-1}\right)$ симметричен, т. е. $m_{il}^{-1} = m_{li}^{-1}$.

Аналогично тому, как вектор, т. е. тензор 1-го ранга, может быть изображен направленным отрезком (стрелкой), симметричный тензор 2-го ранга (T_{ik}) может быть изображен поверхностью второго порядка

$$\sum_{i,\,k} T_{ik} x_i x_k = 1 \qquad \text{(П.11.12)}$$

или, подробнее,

$$T_{11}x_1^2 + T_{22}x_2^2 + T_{33}x_3^2 + 2T_{12}x_1x_2 + 2T_{13}x_1x_3 + 2T_{23}x_2x_3 = 1. \qquad \text{(П.11.12а)}$$

Можно показать, что если все $T_{ii} > 0$, то эта поверхность—эллипсоид, называемый *тензорным эллипсоидом*.

4. Компоненты тензора меняются при переходе от одной координатной системы к другой. Можно доказать, что для симметричного тензора 2-го ранга всегда можно выбрать такую координатную систему, чтобы тензор в ней принял диагональную форму, т. е. были отличны от нуля только компоненты тензора T_{ii}, расположенные вдоль главной диагонали. В этом случае тензорный эллипсоид принимает вид

$$T_{11}x_1^2 + T_{22}x_2^2 + T_{33}x_3^2 = 1. \qquad \text{(П.11.12б)}$$

Мы видим, что с точки зрения аналитической геометрии выбор такой координатной системы соответствует приведению тензорного эллипсоида к главным осям.

Тензоры электропроводности и обратных эффективных масс имеют в главных осях вид

$$(\sigma_{ik}) = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}, \quad \left(m_{ik}^{-1}\right) = \begin{pmatrix} m_1^{-1} & 0 & 0 \\ 0 & m_2^{-1} & 0 \\ 0 & 0 & m_3^{-1} \end{pmatrix}. \qquad \text{(П.11.13)}$$

Величина 1: m_i^{-1}, обозначаемая нами как m_i, не является компонентой тензора, т. е. не преобразуется по закону (П.11.3); она имеет размерность массы и называется эффективной массой электрона в кристалле.

Так как тензоры электропроводности и обратной эффективной массы являются материальными константами кристалла, то направления их главных осей определенным образом ориентированы относительно осей симметрии кристалла (например, в случае кристалла ромбической сингонии главные оси параллельны ребрам элементарной ячейки). В общем случае главные оси разных материальных тензоров, например диэлектрической постоянной и обратных эффективных масс, могут не совпадать. Ответ на этот вопрос может дать только микроскопическая теория этих тензоров. Если кристалл обладает такой симметрией, что главные оси тензора x_1 и x_2 физически эквивалентны, то, очевидно, $\sigma_1=\sigma_2$ и $m_1=m_2$. В кубическом кристалле эквивалентны все три главные оси тензора (параллельные ребрам куба), поэтому, например, тензор электропроводности

$$(\sigma_{ik})=\begin{pmatrix}\sigma & 0 & 0\\ 0 & \sigma & 0\\ 0 & 0 & \sigma\end{pmatrix} \qquad (\text{П.11.13а})$$

и, следовательно, закон (П.11.7) переходит в изотропное соотношение (П.11.6а). Можно сказать, что тензор 2-го ранга в кристаллах кубической симметрии вырождается в скаляр.

Приложение 12

Блоховская функция электрона $\psi_k(\boldsymbol{r})=u_k(\boldsymbol{r})\,e^{i\boldsymbol{kr}}$ в периодическом поле кристалла $V(\boldsymbol{r})$ удовлетворяет уравнению Шредингера

$$\nabla^2\psi(\boldsymbol{r})+\frac{2m}{\hbar^2}[\varepsilon(\boldsymbol{k})-V(\boldsymbol{r})]\,\psi(\boldsymbol{r})=0. \qquad (\text{П.12.1})$$

Дифференцируя обе части уравнения по k_x, получим

$$\nabla^2\frac{\partial\psi}{\partial k_x}+\frac{2m}{\hbar^2}\left[\frac{\partial\varepsilon}{\partial k_x}\psi-(\varepsilon-V)\frac{\partial\psi}{\partial k_x}\right]=0. \qquad (\text{П.12.2})$$

Имеем

$$\frac{\partial\psi}{\partial k_x}=\frac{\partial}{\partial k_x}(ue^{i\boldsymbol{kr}})=ix\psi+e^{i\boldsymbol{kr}}\frac{\partial u}{\partial k_x} \qquad (\text{П.12.3})$$

и

$$\nabla^2\frac{\partial\psi}{\partial k_x}=\nabla^2(ix\psi)+\nabla^2\left(e^{i\boldsymbol{kr}}\frac{\partial u}{\partial k_x}\right)=2i\frac{\partial\psi}{\partial x}+ix\,\nabla^2\psi+\nabla^2\left(e^{i\boldsymbol{kr}}\frac{\partial u}{\partial k_x}\right). \qquad (\text{П.12.4})$$

Подставляя (П.12.4) и (П.12.3) в (П.12.2), получим

$$2i\frac{\partial\psi}{\partial x}+\frac{2m}{\hbar^2}\frac{\partial\varepsilon}{\partial k_x}\psi+ix\left\{\nabla^2\psi+\frac{2m}{\hbar^2}(\varepsilon-V)\psi\right\}+$$
$$+\left[\nabla^2+\frac{2m}{\hbar^2}(\varepsilon-V)\right]e^{i\boldsymbol{kr}}\frac{\partial u}{\partial k_x}=0, \qquad (\text{П.12.5})$$

где фигурная скобка равна нулю в силу (П.12.1).

Умножая обе части уравнения (П.12.5) слева на ψ^* и интегрируя по основной области кристалла, получим

$$2i\int\psi^*\frac{\partial\psi}{\partial x}d\tau+\frac{2m}{\hbar^2}\frac{\partial\varepsilon}{\partial k_x}\int\psi^*\psi\,d\tau+$$
$$+\int e^{i\boldsymbol{kr}}\frac{\partial u}{\partial k_x}\left[\nabla^2+\frac{2m}{\hbar^2}(\varepsilon-V)\right]\psi^*d\tau=0, \qquad (\text{П.12.6})$$

где при преобразовании последнего интеграла мы воспользовались самосопряженностью оператора ∇^2. Последний интеграл равен нулю, так как ψ^* тоже удовлетворяет уравнению (П.12.1). Таким образом, если учесть, что ψ нормирована на основной объем, то из (П.12.6) следует

$$\frac{\hbar}{mi}\int \psi^* \frac{\partial \psi}{\partial x}\, d\tau = \frac{1}{\hbar}\frac{\partial \varepsilon}{\partial k_x}. \qquad (П.12.7)$$

Сравнивая это выражение с (3.28), получим (3.32).

Приложение 13

На рис. П.3 изображен куб и прямоугольная координатная система с осями x, y и z, направленными по его ребрам. Темный кружок ● изображает атом в центре куба, а светлые кружки ○ — атомы на его гранях. Если a — длина ребра куба и $\boldsymbol{i}_0$, $\boldsymbol{j}_0$, $\boldsymbol{k}_0$ — орты координатной системы, то основные векторы решетки гранецентрированного куба могут быть выбраны следующим образом:

$$\boldsymbol{a}_1 = \frac{a}{2}(\boldsymbol{i}_0 + \boldsymbol{k}_0),$$
$$\boldsymbol{a}_2 = \frac{a}{2}(\boldsymbol{i}_0 + \boldsymbol{j}_0), \qquad (П.13.1)$$
$$\boldsymbol{a}_3 = \frac{a}{2}(\boldsymbol{j}_0 + \boldsymbol{k}_0).$$

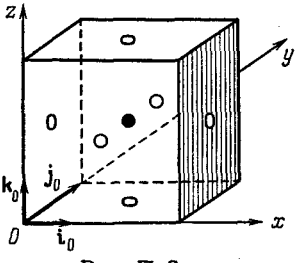


Рис. П. 3.

Векторы $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ и $\boldsymbol{a}_3$ проведены из начала координат к ближайшим ○-атомам.

Из (П.13.1) непосредственно

$$[\boldsymbol{a}_2\boldsymbol{a}_3] = \frac{a^2}{4}(\boldsymbol{i}_0 - \boldsymbol{j}_0 + \boldsymbol{k}_0) \qquad (П.13.1а)$$

и, следовательно, объем элементарной ячейки

$$\Omega_0 = (\boldsymbol{a}_1[\boldsymbol{a}_2\boldsymbol{a}_3]) = a^3/4, \qquad (П.13.1б)$$

т. е. составляет $^1/_4$ объем всего куба.

Для решетки объемноцентрированного куба основные векторы удобно выбрать следующим образом:

$$\boldsymbol{a}_1' = \frac{a}{2}(\boldsymbol{i}_0 - \boldsymbol{j}_0 + \boldsymbol{k}_0), \quad \boldsymbol{a}_2' = \frac{a}{2}(\boldsymbol{i}_0 + \boldsymbol{j}_0 - \boldsymbol{k}_0), \quad \boldsymbol{a}_3' = \frac{a}{2}(-\boldsymbol{i}_0 + \boldsymbol{j}_0 + \boldsymbol{k}_0), \qquad (П.13.2)$$

т. е. направить их из начала координат O к центральным ●-атомам кубов, расположенных спереди, снизу и слева от изображенного на рис. П.3. Аналогично были выбраны основные векторы на рис. I.6, *б*.

Из (П.13.2)

$$[\boldsymbol{a}_2'\boldsymbol{a}_3'] = \frac{a^2}{2}(\boldsymbol{i}_0 + \boldsymbol{k}_0). \qquad (П.13.2а)$$

Объем элементарной ячейки

$$\Omega_0' = (\boldsymbol{a}_1'[\boldsymbol{a}_2'\boldsymbol{a}_3']) = {}^1/_2 a^3, \qquad (П.13.2б)$$

т. е. равен $^1/_2$ объема куба.

Нетрудно теперь для обоих случаев построить обратные решетки. Для гранецентрированной кубической решетки из (П.13.1а) и (П.13.1б)

$$\boldsymbol{b}_1 = \frac{1}{\Omega_0}[\boldsymbol{a}_2\boldsymbol{a}_3] = \frac{1}{a}[\boldsymbol{i}_0 - \boldsymbol{j}_0 + \boldsymbol{k}_0]$$

и аналогично

$$\boldsymbol{b}_2=\frac{1}{\Omega_0}[\boldsymbol{a}_3\boldsymbol{a}_1]=\frac{1}{a}(\boldsymbol{i}_0+\boldsymbol{j}_0-\boldsymbol{k}_0),$$
$$\boldsymbol{b}_3=\frac{1}{\Omega_0}[\boldsymbol{a}_1\boldsymbol{a}_2]=\frac{1}{a}(-\boldsymbol{i}_0+\boldsymbol{j}_0+\boldsymbol{k}_0). \qquad (\text{П.13.3})$$

Для решетки объемноцентрированного куба аналогично из (П.13.2а) и (П.13.2б)

$$\boldsymbol{b}_1'=\frac{1}{\Omega_0'}[\boldsymbol{a}_2'\boldsymbol{a}_3']=\frac{1}{a}(\boldsymbol{i}_0+\boldsymbol{k}_0),$$
$$\boldsymbol{b}_2'=\frac{1}{\Omega_0'}[\boldsymbol{a}_3'\boldsymbol{a}_1']=\frac{1}{a}(\boldsymbol{i}_0+\boldsymbol{j}_0), \qquad (\text{П.13.4})$$
$$\boldsymbol{b}_3'=\frac{1}{\Omega_0'}[\boldsymbol{a}_1'\boldsymbol{a}_2']=\frac{1}{a}(\boldsymbol{j}_0+\boldsymbol{k}_0).$$

Сравнивая (П.13.3) с (П.13.2) и (П.13.4) с (П.13.1), видим, что для гранецентрированного куба обратная решетки — объемно-центрированный куб и наоборот.

Приложение 14

Для вычисления суммы $\sum_{n_0}^{1,8} e^{i\boldsymbol{k}\boldsymbol{a}_{n_0}}$ в случае решетки объемноцентрированного куба обратимся к рис. П.3. Вектор $\boldsymbol{a}_{n_0}=\boldsymbol{a}_1$. Для ●-атома в центре куба

$$\boldsymbol{a}_1=\frac{a}{2}(\boldsymbol{i}_0+\boldsymbol{j}_0+\boldsymbol{k}_0). \qquad (\text{П.14.1})$$

Аналогично для остальных семи атомов

$$\boldsymbol{a}_2=\frac{a}{2}(-\boldsymbol{i}_0+\boldsymbol{j}_0+\boldsymbol{k}_0),\quad \boldsymbol{a}_3=\frac{a}{2}(-\boldsymbol{i}_0-\boldsymbol{j}_0+\boldsymbol{k}_0),$$
$$\boldsymbol{a}_4=\frac{a}{2}(\boldsymbol{i}_0-\boldsymbol{j}_0+\boldsymbol{k}_0),\quad \boldsymbol{a}_5=\frac{a}{2}(\boldsymbol{i}_0+\boldsymbol{j}_0-\boldsymbol{k}_0), \qquad (\text{П.14.1а})$$
$$\boldsymbol{a}_6=\frac{a}{2}(-\boldsymbol{i}_0+\boldsymbol{j}_0-\boldsymbol{k}_0),\quad \boldsymbol{a}_7=\frac{a}{2}(-\boldsymbol{i}_0-\boldsymbol{j}_0-\boldsymbol{k}_0),$$
$$\boldsymbol{a}_8=\frac{a}{2}(\boldsymbol{i}_0-\boldsymbol{j}_0-\boldsymbol{k}_0).$$

Из (П.14.1) следует

$$\boldsymbol{k}\boldsymbol{a}_1=\frac{a}{2}(k_x+k_y+k_z)$$

и аналогично для произведений $\boldsymbol{k}\boldsymbol{a}_i$ ($i=2, 3, \ldots, 8$) из (П.14.1а).

Таким образом,

$$\sum_{n_0}^{1,8} e^{i\boldsymbol{k}\boldsymbol{a}_{n_0}}=e^{\frac{ia}{2}(k_x+k_y+k_z)}+e^{\frac{ia}{2}(-k_x+k_y+k_z)}+$$
$$+e^{\frac{ia}{2}(-k_x-k_y+k_z)}+e^{\frac{ia}{2}(k_x-k_y+k_z)}+e^{\frac{ia}{2}(k_x+k_y-k_z)}+$$
$$+e^{\frac{ia}{2}(-k_x+k_y-k_z)}+e^{\frac{ia}{2}(-k_x-k_y-k_z)}+e^{\frac{ia}{2}(k_x-k_y-k_z)}.$$

Вынося из 1-го и 5-го слагаемых множитель $e^{\frac{ia}{2}(k_x+k_y)}$, получим для них

$$e^{\frac{ia}{2}(k_x+k_y)}\left[e^{\frac{iak_z}{2}}+e^{-\frac{iak_z}{2}}\right]=2e^{\frac{ia}{2}(k_x+k_y)}\cos\frac{ak_z}{2}.$$

Производя аналогичную операцию для 2-го и 6-го слагаемых, 3-го и 7-го слагаемых и 4-го и 8-го слагаемых, получим в каждом случае выражение, содержащее один и тот же множитель $2\cos\frac{ak_z}{2}$. Вынося его за общую скобку, преобразуем оставшееся выражение аналогично. В результате придем к формуле (7.10б).

Приложение 15

Подставим волновую функцию электрона в кристалле (7.17) в соответствующее уравнение Шредингера (7.3)

$$\sum_n e^{i\boldsymbol{k}\boldsymbol{a}_n}\left[-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r})-\varepsilon\right]\times$$
$$\times\{\alpha\psi_x(\boldsymbol{r}-\boldsymbol{a}_n)+\beta\psi_y(\boldsymbol{r}-\boldsymbol{a}_n)+\gamma\psi_z(\boldsymbol{r}-\boldsymbol{a}_n)\}=0. \quad (\text{П}.15.1)$$

Так как ψ_x, ψ_y, ψ_z — волновые функции электрона в изолированном атоме, то они удовлетворяют уравнению

$$-\frac{\hbar^2}{2m}\nabla^2\psi_\mu(\boldsymbol{r}-\boldsymbol{a}_n)+[\mathcal{U}(|\boldsymbol{r}-\boldsymbol{a}_n|)-\varepsilon_0]\psi_\mu(\boldsymbol{r}-\boldsymbol{a}_n)=0, \quad (\text{П}.15.2)$$

где $\mu=x$, y или z, ε_0 — энергия p-электрона, а $\mathcal{U}(|\boldsymbol{r}-\boldsymbol{a}_n|)$ его потенциальная энергия в поле изолированного $\boldsymbol{n}$-го узла.

Заменяя в (П.15.1) величины $-\frac{\hbar^2}{2m}\nabla^2\psi_\mu$, согласно уравнению (П.15.2), получим

$$\sum_n e^{i\boldsymbol{k}\boldsymbol{a}_n}(\varepsilon-\varepsilon_0)[\alpha\psi_x(\boldsymbol{r}-\boldsymbol{a}_n)+\beta\psi_y(\boldsymbol{r}-\boldsymbol{a}_n)+\gamma\psi_z(\boldsymbol{r}-\boldsymbol{a}_n)]=$$
$$=\sum_n e^{i\boldsymbol{k}\boldsymbol{a}_n}[V(\boldsymbol{r})-\mathcal{U}(|\boldsymbol{r}-\boldsymbol{a}_n|)]\times$$
$$\times\{\alpha\psi_x(\boldsymbol{r}-\boldsymbol{a}_n)+\beta\psi_y(\boldsymbol{r}-\boldsymbol{a}_n)+\gamma\psi_z(\boldsymbol{r}-\boldsymbol{a}_n)\}. \quad (\text{П}.15.3)$$

Умножим обе части этого равенства на волновую функцию нулевого узла $\psi_x(\boldsymbol{r})$ [1] и проинтегрируем по объему основной области кристалла. Предполагая волновые функции ψ_μ ортонормированными, т. е.

$$\int\psi_\mu(\boldsymbol{r})\psi_\nu(\boldsymbol{r})\,d\tau=\delta_{\mu\nu}\quad(\mu\text{ и }\nu=x,\ y,\ z), \quad (\text{П}.15.4)$$

и пренебрегая интегралом перекрытия волновых функций разных узлов, получим для левой части (П.15.3)

$$\alpha(\varepsilon-\varepsilon_0).$$

---

[1]) Очевидно, что по отношению к сумме $\sum_n$, взятой по всем узлам кристалла в (П.15.3), нулевой узел эквивалентен любому узлу решетки $\boldsymbol{n}'$; можно было бы умножить (П.15.3) на $\psi(\boldsymbol{r}-\boldsymbol{a}_{n'})$.

Правая часть (П.15.3) для $\boldsymbol{n}=0$ будет содержать интегралы вида

$$\mathscr{T}_{x\nu}=\int \psi_x(\boldsymbol{r})\,[V(\boldsymbol{r})-\mathcal{U}(\boldsymbol{r})]\,\psi_\nu(\boldsymbol{r})\,d\tau \quad (\nu=x, y, z). \tag{П.15.6}$$

Для $\nu=x$ интеграл

$$\mathscr{T}_{xx}=\int \psi_x^2(\boldsymbol{r})\,[V(\boldsymbol{r})-\mathcal{U}(\boldsymbol{r})]\,d\tau=-C<0 \tag{П.15.7}$$

аналогично выражению (7.7). Для $\nu=y$ или $\nu=z$

$$\mathscr{T}_{xy}=\mathscr{T}_{xz}=0. \tag{П.15.8}$$

В последнем можно убедиться, повернув координатную систему вокруг оси z так, чтобы ось x совпала с осью y, а ось y — с осью $-x$. При этом, согласно (7.16), функция ψ_x превратится в ψ_y, а ψ_y — в $-\psi_x$. Так как $[V(\boldsymbol{r})-\mathcal{U}(\boldsymbol{r})]$ при таком преобразовании координат в силу кубической симметрии поля кристалла останется без изменения, то весь интеграл изменит знак на противоположный. С другой стороны, при любом преобразовании координат величина определенного интеграла остается неизменной; таким образом, $\mathscr{T}_{xy}=-\mathscr{T}_{xy}$, откуда непосредственно следует (П.15.8). Из тех же соображений кубической симметрии поля кристалла следует, что

$$\mathscr{T}_{xx}=\mathscr{T}_{yy}=\mathscr{T}_{zz}=-C. \tag{П.15.7а}$$

Учтем в правой части (П.15.3) ближайших соседей нулевого узла, т. е. члены с $\boldsymbol{a}_n=\pm a\boldsymbol{i}_0,\ \pm a\boldsymbol{j}_0,\ \pm a\boldsymbol{k}_0$, где $\boldsymbol{i}_0$, $\boldsymbol{j}_0$ и $\boldsymbol{k}_0$ — орты прямоугольных осей x, y и z. Им соответствуют интегралы вида

$$\mathscr{T}_{x\nu}(\boldsymbol{a}_{n_0})=\int \psi_x(\boldsymbol{r})\,[V(\boldsymbol{r})-\mathcal{U}(|\boldsymbol{r}-\boldsymbol{a}_{n_0}|)]\,\psi_\nu(\boldsymbol{r}-\boldsymbol{a}_{n_0})\,d\tau, \tag{П.15.9}$$

где $\nu=x$, y или z.

Рассмотрим вначале случай $\nu=y$ или z. Возьмем соседний атом $\boldsymbol{n}_0$ по оси z [1]), т. е. положим $\boldsymbol{a}_{n_0}=a\boldsymbol{k}_0$. Если учитывать только поле, создаваемое нулевым узлом и выбранным нами соседним атомом по оси z, а это законно в той области, где волновые функции в (П.15.9) заметно отличны от нуля, то $[V(\boldsymbol{r})-\mathcal{U}(|\boldsymbol{r}-a\boldsymbol{k}_0|)]$ от азимутального угла φ не зависит. Так как $\psi_x\sim\cos\varphi$, $\psi_y\sim\sin\varphi$ и ψ_z от φ не зависит, то легко видеть, что интеграл (П.15.9) при интегрировании по φ в случаях $\nu=y$ или $\nu=z$ равен нулю. Следует, однако, отметить, что в случае $\nu=x$, когда интеграл (П.15.9) отличен от нуля, он будет иметь разное значение в зависимости от того, расположен ли соседний атом вдоль оси x или вдоль осей y или z. В первом случае волновые функции $\psi_x(\boldsymbol{r})$ и $\psi_x(\boldsymbol{r}-a\boldsymbol{i}_0)$, меняющие свой знак при прохождении плоскостей, перпендикулярных к оси x, так, как это изображено на рис. П.4, имеют разные знаки в области их перекрытия, и так как $[V(\boldsymbol{r})-\mathcal{U}(|\boldsymbol{r}\pm a\boldsymbol{i}_0|)]<0$, то можно думать, что интеграл (П.15.9) в этом случае положителен, т. е.

Рис. П. 4.

$$\mathscr{T}_{xx}(\pm ai_0)=\int \psi_x(\boldsymbol{r})\,[V(\boldsymbol{r})-\mathcal{U}(|\boldsymbol{r}\mp a\boldsymbol{i}_0|)]\,\psi_x(\boldsymbol{r}\mp a\boldsymbol{i}_0)\,d\tau=A>0. \tag{П.15.10}$$

Во втором случае интеграл будет иметь иное значение, и сказать что-либо

[1]) Мы выбираем соседний атом по оси z только для того, чтобы удобно использовать привычно ориентированную систему сферических координат.

о его знаке еще более затруднительно. Мы обозначим

$$\mathscr{T}_{xx}(\pm aj_0)=\mathscr{T}_{xx}(\pm ak_0)=$$
$$=\int \psi_x(\boldsymbol{r})\,[V(\boldsymbol{r})-\mathcal{U}(|\boldsymbol{r}\mp aj_0|)]\,\psi_x(\boldsymbol{r}\mp aj_0)\,d\tau=-B. \quad (\text{П.15.11})$$

Учитывая шесть ближайших соседей простой кубической решетки, получим для правой части (П.15.3)

$$\alpha\left[-C+A\left(e^{ik_xa}+e^{-ik_xa}\right)-B\left(e^{ik_ya}+e^{-ik_ya}+e^{ik_za}+e^{-ik_za}\right)\right]=$$
$$=\alpha\left[-C+2A\cos ak_x-2B\left(\cos ak_y+\cos ak_z\right)\right]. \quad (\text{П.15.12})$$

Из (П.15.3), (П.15.5) и (П.15.12) следует

$$\alpha\left[\varepsilon-\varepsilon_0+C-2A\cos ak_x+2B\left(\cos ak_y+\cos ak_z\right)\right]=0. \quad (\text{П.15.13})$$

Если мы умножим (П.15.3) на $\psi_y(\boldsymbol{r})$ или $\psi_z(\boldsymbol{r})$ и проинтегрируем по объему основной области кристалла, то получим вместо (П.15.13) два других равенства, которые могут быть получены при циклической перестановке x, y и z и α, β и γ, т. е.

$$\beta\left[\varepsilon-\varepsilon_0+C-2A\cos ak_y+2B\left(\cos ak_z+\cos ak_x\right)\right]=0, \quad (\text{П.15.13а})$$

$$\gamma\left[\varepsilon-\varepsilon_0+C-2A\cos ak_z+2B\left(\cos ak_x+\cos ak_y\right)\right]=0. \quad (\text{П.15.13б})$$

Равенства (П.15.13), (П.15.13а), (П.15.13б) могут быть одновременно удовлетворены тривиальным образом, если положить $\alpha=\beta=\gamma=0$, что не представляет интереса, так как при этом волновая функция (7.17) тождественно равна нулю. Если $\alpha\neq 0$, то из (П.15.13) следует

$$\varepsilon=\varepsilon_0-C+2A\cos ak_x-2B\left(\cos ak_y+\cos ak_z\right) \quad (\text{П.15.14})$$

и, следовательно, $\beta=\gamma=0$, так как в противном случае ε будет зависеть от k_x, k_y и k_z не согласно (П.15.14), а так, как это будет следовать из (П.15.13а), (П.15.13б) при $\beta\neq 0$ и $\gamma\neq 0$. Таким образом, мы видим, что возможны три случая: 1) $\alpha\neq 0$, $\beta=\gamma=0$; 2) $\beta\neq 0$, $\alpha=\gamma=0$ и 3) $\gamma\neq 0$, $\alpha=\beta=0$, соответствующие законам дисперсии (7.18), (7.18а), (7.18б).

Приложение 16

Обозначим проекции на ось x векторов $\boldsymbol{r}_n$, $\boldsymbol{r}_p$, $\boldsymbol{R}$ и $\boldsymbol{r}$ теми же, но нежирными буквами, так как (3.2) имеют место и для проекций, то они справедливы для $r_n r_p R$ и r.

Вычислим $\partial^2\psi/\partial r_n^2$

$$\frac{\partial\psi}{\partial r_n}=\frac{\partial\psi}{\partial R}\frac{\partial R}{\partial r_n}+\frac{\partial\psi}{\partial r}\frac{\partial\rho}{\partial r_n}=\frac{m_n}{M}\frac{\partial\psi}{\partial R}+\frac{\partial\psi}{\partial r},$$

$$\frac{\partial^2\psi}{\partial r_n^2}=\left(\frac{m_n}{M}\frac{\partial^2\psi}{\partial R^2}+\frac{\partial^2\psi}{\partial r\,\partial R}\right)\frac{m_n}{M}+\left(\frac{m_n}{M}\frac{\partial^2\psi}{\partial R\,\partial r}+\frac{\partial^2\psi}{\partial r^2}\right)=$$
$$=\left(\frac{m_n}{M}\right)^2\frac{\partial^2\psi}{\partial R^2}+2\frac{m_n}{M}\frac{\partial^2\psi}{\partial r\,\partial R}+\frac{\partial^2\psi}{\partial r^2}.$$

Аналогично

$$\frac{\partial^2\psi}{\partial r_p^2}=\left(\frac{m_p}{M}\right)^2\frac{\partial^2\psi}{\partial R^2}-2\frac{m_p}{M}\frac{\partial^2\psi}{\partial r\,\partial R}+\frac{\partial^2\psi}{\partial r^2},$$

откуда

$$\frac{1}{m_n}\frac{\partial^2\psi}{\partial r_n^2}+\frac{1}{m_p}\frac{\partial^2\psi}{\partial r_p^2}=\frac{1}{M}\frac{\partial^2\psi}{\partial R^2}+\frac{1}{\mu}\frac{\partial^2\psi}{\partial r^2}.$$

Так как для проекций на оси y и z имеют место аналогичные равенства, то отсюда непосредственно следует (3.3).

Приложение 17

Для вычисления интеграла I (2.7) в следующем приближении введем переменную $\eta=\frac{\varepsilon-\zeta}{k_0T}$ и разложим функцию $\varphi(\varepsilon)=\psi(\eta)$ в ряд по степеням η. В силу дельтаобразного характера функции $\left(-\frac{\partial f_0}{\partial\varepsilon}\right)$ в интеграле играют роль только значения ε, близкие к ζ, т. е. малые η

$$\psi(\eta)=\psi(0)+\psi'(0)\,\eta+\frac{1}{2}\psi''(0)\,\eta^2+\ldots \qquad (\text{П}.17.1)$$

Далее,

$$-\frac{\partial f_0}{\partial\varepsilon}\,d\varepsilon=-\frac{\partial f_0}{\partial\eta}\,d\eta=-\frac{\partial}{\partial\eta}\left(\frac{1}{e^{\eta}+1}\right)d\eta=\frac{e^{-\eta}}{(1+e^{-\eta})^2}\,d\eta. \qquad (\text{П}.17.2)$$

Заменяя нижний предел в интеграле I, равный $-z$ (для переменной η) на $-\infty$, что допустимо, так как $z\gg 1$, получим

$$I=\int\limits_{-\infty}^{+\infty}\psi(\eta)\left(-\frac{\partial f_0}{\partial\eta}\right)d\eta=\psi(0)\int\limits_{-\infty}^{+\infty}\left(-\frac{\partial f_0}{\partial\eta}\right)d\eta+$$
$$+\psi'(0)\int\limits_{-\infty}^{+\infty}\frac{\eta e^{-\eta}}{(1+e^{-\eta})^2}\,d\eta+\frac{1}{2}\psi''(0)\int\limits_{-\infty}^{+\infty}\frac{\eta^2e^{-\eta}}{(1+e^{-\eta})^2}\,d\eta. \qquad (\text{П}.17.3)$$

Первый интеграл правой части равен единице, второй — в силу нечетности подынтегральной функции — нулю, третий равен

$$\frac{1}{2}\int\limits_{-\infty}^{+\infty}\frac{\eta^2e^{-\eta}}{(1+e^{-\eta})^2}\,d\eta=\int\limits_{0}^{\infty}\frac{\eta^2e^{-\eta}}{(1+e^{-\eta})^2}\,d\eta=\int\limits_{0}^{\infty}\eta^2\left[e^{-\eta}-2e^{-2\eta}+3e^{-3\eta}-\ldots\right]d\eta=$$
$$=2\left(1-\frac{1}{2^2}+\frac{1}{3^2}-\frac{1}{4^2}+\ldots\right)=2\,\frac{\pi^2}{12}=\frac{\pi^2}{6}. \qquad (\text{П}.17.4)$$

При вычислении последнего интеграла мы разложили множитель $(1+e^{-\eta})^{-2}$ подынтегральной функции в ряд по $e^{-\eta}$, а в конце воспользовались формулой для суммы знакопеременного ряда обратных значений квадратов натурального ряда чисел [1]).

Окончательно

$$I=\psi(0)+\frac{\pi^2}{6}\psi''(0). \qquad (\text{П}.17.5)$$

В случае (2.5)

$$\psi(\eta)=\frac{\sqrt{2}}{\pi^2}\,\frac{m^{*3/2}}{\hbar^3}\,\frac{2}{3}\,\zeta+k_0T\eta^{3/2}. \qquad (\text{П}.17.6)$$

Используя (П.17.5) и (П.17.6), получим для (2.5)

$$n=\frac{(2m^*)^{3/2}}{3\pi^2\hbar^3}\,\zeta^{3/2}\left[1+\frac{\pi^2}{8}\left(\frac{k_0T}{\zeta}\right)^2\right].$$

Считая, что в поправочном члене в квадратных скобках можно положить $\zeta=\zeta_0$, и решая полученное равенство относительно ζ, которое стоит перед

[1]) Смирнов В. И., т. II, п. 156.

квадратной скобкой, получим

$$\zeta=\frac{\hbar^2}{2m^*}(3\pi^2 n)^{2/3}\left[1+\frac{\pi^2}{8}\left(\frac{k_0T}{\zeta_0}\right)^2\right]^{-2/3}.$$

Разлагая квадратную скобку в ряд по $(k_0T/\zeta_0)^2$ и ограничиваясь первыми двумя членами, получим формулу (2.9) текста.

Приложение 18

По формуле Пуассона [1])

$$\sum_{N=-\infty}^{\infty}\varphi(2\pi N+t)=\frac{1}{2\pi}\sum_{l=-\infty}^{\infty}e^{ilt}\int_{-\infty}^{\infty}\varphi(\tau)\,e^{-il\tau}\,d\tau. \qquad (\text{П}.18.1)$$

В нашем случае (5.25б)

$$\varphi(2\pi N+t)=\frac{1}{(2\pi)^{3/2}}(2\pi N+t)^{3/2}, \qquad (\text{П}.18.2)$$

если заменить индекс суммирования N на $-N$ и положить

$$t=2\pi\varepsilon-\pi. \qquad (\text{П}.18.3)$$

Так как $\varphi(2\pi N+t)$ — вещественно, то в нашем случае

$$-[\varepsilon-1/2]\leqslant N\leqslant 0, \qquad (\text{П}.18.4)$$

где $[x]$ обозначает наибольшее целое число, заключенное в величине x. Пределы изменения аргумента функции $\varphi(2\pi N+t)$ от $2\pi(\varepsilon-1/2)-2\pi[\varepsilon-1/2]$ (что при больших ε, как мы и будем предполагать дальше, близко к нулю) — до t. Это определяет пределы интеграла в (П.18.1), так как вне этих пределов $\varphi(\tau)\equiv 0$.

Таким образом, $\Phi(\varepsilon)$ в (5.25а) для больших ε равно

$$\Phi(\varepsilon)=\frac{1}{(2\pi)^{5/2}}\sum_{l=-\infty}^{\infty}e^{ilt}\int_0^t\tau^{3/2}e^{-il\tau}\,d\tau. \qquad (\text{П}.18.5)$$

Слагаемое для $l=0$ равно

$$\int_0^t\tau^{3/2}\,d\tau=\frac{2}{5}\,t^{5/2}. \qquad (\text{П}.18.6)$$

Для суммы двух слагаемых $+l$ и $-l$ получим

$$(-1)^l\left\{\left[\cos(2\pi l\varepsilon)+i\sin(2\pi l\varepsilon)\right]\int_0^t\tau^{3/2}\left[\cos(l\tau)-i\sin(l\tau)\right]d\tau+\right.$$
$$\left.+\left[\cos(2\pi l\varepsilon)-i\sin(2\pi l\varepsilon)\right]\int_0^t\tau^{3/2}\left[\cos(l\tau)+i\sin(l\tau)\right]d\tau\right\}=$$
$$=(-1)^l\,2\left[\cos(2\pi l\varepsilon)\int_0^t\tau^{3/2}\cos(l\tau)\,d\tau+\sin(2\pi l\varepsilon)\int_0^t\tau^{3/2}\sin(l\tau)\,d\tau\right]. \qquad (\text{П}.18.7)$$

---

[1]) Курант Р., Гильберт Д. Методы математической физики,— 2-е изд.—М.—Л., 1951, т. 1, с. 71.

При этом мы учли, что $e^{\pm il\pi}=(-1)^l$, и воспользовались соотношением $e^{\pm i\alpha}=\cos\alpha\pm i\sin\alpha$. Интегрируя в (П.18.7) по частям, так чтобы понизить степень у $\tau^{3/2}$ до $\tau^{-1/2}$, и заменяя в последнем интеграле переменную τ на $(\pi/2l)x^2$, получим

$$\Phi(\varepsilon)=\frac{1}{(2\pi)^{5/2}}\left[\frac{2}{5}(2\pi\varepsilon)^{5/2}+3\sum_{l=1}^{\infty}\frac{(-1)^l}{l^{5/2}}\left\{(2\pi l\varepsilon)^{1/2}-\right.\right.$$

$$\left.\left.-\sqrt{\frac{\pi}{2}}\left[\sin(2\pi l\varepsilon)\,S\left(\sqrt{4l\varepsilon}\right)+\cos(2\pi l\varepsilon)\,C\left(\sqrt{4l\varepsilon}\right)\right]\right\}\right], \qquad (\text{П.18.8})$$

где

$$S(u)=\int_0^u \sin\left(\frac{\pi}{2}x^2\right)dx, \qquad (\text{П.18.8а})$$

$$C(u)=\int_0^u \cos\left(\frac{\pi}{2}x^2\right)dx \qquad (\text{П.18.8б})$$

— интегралы Френеля [1]); функции $S(u)$ и $C(u)$, так же как и тригонометрические функции, входящие в (П.18.8) с $l=1$, осциллируют с периодом порядка 1, но с затухающей амплитудой; $C(\infty)=S(\infty)=0{,}5$.

Сумма, входящая в (П.18.8), равна [2])

$$\sum_{l=1}^{\infty}\frac{(-1)^l}{l^{5/2}}(2\pi l\varepsilon)^{1/2}=(2\pi\varepsilon)^{1/2}\sum_{l=1}^{\infty}\frac{(-1)^l}{l^2}=(2\pi\varepsilon)^{1/2}\left(-\frac{\pi^2}{12}\right). \qquad (\text{П.18.9})$$

Из (П.18.8) и (П.18.9) сразу следует (5.26) основного текста.

Приложение 19

Для вывода соотношений Крамерса—Кронига (1.9а) продолжим аналитически функцию $\varepsilon(\omega)$ на верхнюю часть комплексной плоскости, положив для этого

$$\omega=\omega_1+i\omega_2 \quad (\omega_2>0). \qquad (\text{П.19.1})$$

Тогда (1.7а) приобретает вид

$$\varepsilon(\omega)-1=\varepsilon(\omega_1+i\omega_2)-1=\int_0^{\infty} f(t)\,e^{i\omega_1 t}e^{-i\omega_2 t}\,dt. \qquad (\text{П.19.2})$$

Из свойств функции $f(t)$ и наличия под знаком интеграла убывающей экспоненты $\exp(-\omega_2 t)$ следует, что интеграл сходится и, следовательно, функция $\varepsilon(\omega)-1$ не имеет в верхней полуплоскости особых точек.

Если функция $\chi(\omega)=\chi(\omega_1+i\omega_2)$ не имеет особых точек внутри замкнутого контура C, то по теореме Коши [3])

$$\oint_C \chi(\omega)\,d\omega=0. \qquad (\text{П.19.3})$$

---

[1]) Лебедев Н. Н. Специальные функции и их приложения.— 2-е изд.— М.— Л., 1963, § 2, п. 4.

[2]) См. Смирнов В. И., т. 2, п. 156.

[3]) Смирнов В. И., т. III, ч. 2, п. 5.

Положим

$$\chi(\omega)=\frac{\varepsilon(\omega)-1}{\omega-\omega_0}, \qquad (\text{П.19.4})$$

где ω_0 — некоторое фиксированное положительное вещественное значение ω. Выберем контур C, как показано на рис. П.5, т. е. в форме полуокружности большого радиуса R, двух отрезков по вещественной оси и полуокружности малого радиуса ρ с центром в точке ω_0; в этом случае интеграл (П.19.3) со значением $\chi(\omega)$ (П.19.4) равен

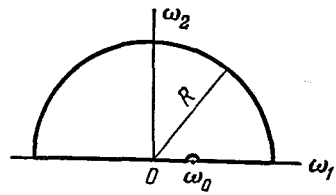


Рис. П. 5.

$$\int\limits_{-R}^{\omega_0-\rho}\frac{\varepsilon(\omega_1)-1}{\omega_1-\omega_0}\,d\omega_1+\int\limits_{(\rho)}\frac{\varepsilon(\omega)-1}{\omega-\omega_0}\,d\omega+$$

$$+\int\limits_{\omega_0+\rho}^{R}\frac{\varepsilon(\omega_1)-1}{\omega_1-\omega_0}\,d\omega_1+\int\limits_{(R)}\frac{\varepsilon(\omega)-1}{\omega-\omega_0}\,d\omega=0. \qquad (\text{П.19.5})$$

Перейдем к пределу $R\longrightarrow\infty$ и $\rho\longrightarrow 0$; последний интеграл равен нулю, так как $\varepsilon(\omega)-1$ экспоненциально стремится к нулю при $R\longrightarrow\infty$. Для вычисления второго интеграла учтем, что $\omega-\omega_0=\rho e^{i\varphi}$, так что при заданном ρ $d\omega=i\rho e^{i\varphi}\,d\varphi$, тогда при $\rho\longrightarrow 0$

$$\int\limits_{\rho\to 0}\frac{\varepsilon(\omega)-1}{\omega-\omega_0}\,d\omega=\int\limits_{\rho\to 0}\frac{\varepsilon(\omega)-1}{\rho e^{i\varphi}}\,i\rho e^{i\varphi}\,d\varphi=i\,[\varepsilon(\omega_0)-1]\,(-\pi). \qquad (\text{П.19.5а})$$

Наконец, сумма первого и третьего интегралов при $\rho\longrightarrow 0$ дает главное значение интеграла; таким образом, из (П.19.5а) получим

$$\int\limits_{-\infty}^{\infty}\frac{\varepsilon(\omega_1)-1}{\omega_1-\omega_0}\,d\omega_1-i\pi\,[\varepsilon(\omega_0)-1]=0,$$

что может быть переписано в виде

$$\varepsilon_1(\omega)+i\varepsilon_2(\omega)-1=\frac{1}{i\pi}\int\frac{\varepsilon_1(x)+i\varepsilon_2(x)-1}{x-\omega}\,dx, \qquad (\text{П.19.6})$$

где мы обозначили $\omega_0\equiv\omega$ и $\omega_1\equiv x$.

Отделяя в последнем равенстве вещественные и мнимые части, получим

$$\varepsilon_1(\omega)-1=\frac{1}{\pi}\int\limits_{-\infty}^{\infty}\frac{\varepsilon_2(x)}{x-\omega}\,dx, \qquad (\text{П.19.7})$$

$$\varepsilon_2(\omega)=-\frac{1}{\pi}\int\limits_{-\infty}^{\infty}\frac{\varepsilon_1(x)-1}{x-\omega}\,dx, \qquad (\text{П.19.7а})$$

что совпадает с первыми равенствами (1.7) и (1.7а).

Для того чтобы получить второе равенство в (1.7), преобразуем интеграл в (П.19.7):

$$\int\limits_{-\infty}^{\infty}\frac{\varepsilon_2(x)}{x-\omega}\,dx=\int\limits_{-\infty}^{0}\frac{\varepsilon_2(x)}{x-\omega}\,dx+\int\limits_{0}^{\infty}\frac{\varepsilon_2(x)}{x-\omega}\,dx.$$

Заменим в первом интеграле правой части x на $-x$ и воспользуемся нечетностью функции $\varepsilon_2(x)$ (1.6а); правая часть равенства приобретает вид

$$\int_0^\infty \frac{\varepsilon_2(x)}{x+\omega}\,dx+\int_0^\infty \frac{\varepsilon_2(x)}{x-\omega}\,dx=2\int_0^\infty \frac{x(\varepsilon_2)\,x}{x^2-\omega^2}\,dx,$$

что совпадает с (1.7).

Аналогично получается второе равенство в (1.7а).

Приложение 20

1. Рассмотрим некоторые существенные для нас особенности теории квантовых переходов.

Пусть полный гамильтониан системы

$$\hat{\mathcal{H}}(t)=\hat{\mathcal{H}}_0+\hat{\mathcal{H}}'(t), \qquad (\text{П.20.1})$$

где $\hat{\mathcal{H}}'(t)$ — малое возмущение, зависящее от времени. Временно́е уравнение Шредингера с гамильтонианом (П.20.1) имеет вид

$$i\hbar\frac{\partial\Psi(t)}{\partial t}=\hat{\mathcal{H}}(t)\,\Psi(t). \qquad (\text{П.20.1а})$$

Разложим решение этого уравнения $\Psi(t)$ по полной системе собственных функций $\psi_n=u_n\exp\left(-\frac{i\varepsilon_n t}{\hbar}\right)$ невозмущенного гамильтониана $\hat{\mathcal{H}}_0$

$$\Psi(t)=\sum_n a_n(t)\,u_n\exp\left(-\frac{i\varepsilon_n t}{\hbar}\right); \qquad (\text{П.20.2})$$

при этом

$$\hat{\mathcal{H}}_0 u_n=\varepsilon_n u_n. \qquad (\text{П.20.2а})$$

Полагая $a_n(t)=a_n^{(0)}+a_n^{(1)}(t)+a_n^{(2)}(t)+\ldots$, где $a_n^{(0)}$ — невозмущенное (начальное) значение $a_n(t)$, а $a_n^{(1)}(t)$, $a_n^{(2)}(t)$ — поправки первого, второго порядка малости по $\hat{\mathcal{H}}'(t)$.

Можно показать, что [1])

$$\frac{da_k^{(1)}}{dt}=\frac{1}{i\hbar}\sum_n \mathcal{H}'_{kn}a_n^{(0)}\exp\frac{i}{\hbar}(\varepsilon_k-\varepsilon_n)\,t, \qquad (\text{П.20.3})$$

$$\frac{da_k^{(2)}}{dt}=\frac{1}{i\hbar}\sum_n \mathcal{H}'_{kn}a_n^{(1)}\exp\frac{i}{\hbar}(\varepsilon_k-\varepsilon_n)\,t, \qquad (\text{П.20.3а})$$

где матричный элемент

$$\mathcal{H}'_{kn}=\int u_k^*\hat{\mathcal{H}}'u_n\,d\tau, \qquad (\text{П.20.4})$$

и $\sum_n$ подразумевает суммирование по дискретным и интегрирование по непрерывным состояниям невозмущенной системы.

2. Пусть в начальный момент времени ($t=0$) система находится в i-м квантовом состоянии, тогда $a_i^{(0)}=1$, а все остальные $a_n^{(0)}=0\,(n\neq i)$. Нас интересует амплитуда $a_f^{(1)}(t)$ конечного состояния f к моменту времени t, если возмущение $\hat{\mathcal{H}}'(t)$ «включается» в момент $t=0$.

---

[1]) Шифф Л. Квантовая механика.— М., 1957, § 29.

Очевидно, что $a_f^{(1)}(0)=0$, поэтому из (П.20.3)

$$a_f^{(1)}(t)=\frac{1}{i\hbar}\int_0^t \mathscr{H}'_{fi}(t')\exp\frac{i}{\hbar}(\varepsilon_f-\varepsilon_i)\,t'\,dt'. \qquad (П.20.5)$$

Если $\mathscr{H}'$ от времени не зависит [1]), то получим

$$a_f^{(1)}(t)=-\frac{1}{\hbar}\mathscr{H}'_{fi}\frac{\exp(i\omega_{fi}t)-1}{\omega_{fi}}, \qquad (П.20.6)$$

где $\mathscr{H}'_{fi}=\text{const}$, а $\omega_{fi}=\dfrac{\varepsilon_f-\varepsilon_i}{\hbar}$.

Из квантовой механики известно, что квадраты модулей коэффициентов в разложении (П.20.2) определяют относительные вероятности соответствующих состояний; поэтому вероятность найти систему в состоянии f в момент времени t, равна

$$\left|a_f^{(1)}(t)\right|^2=\left|\mathscr{H}'_{fi}\right|^2\frac{4\sin^2\left(\frac{\omega_{fi}t}{2}\right)}{\hbar^2\omega_{fi}^2}=\frac{2\pi t}{\hbar^2}\left|\mathscr{H}'_{fi}\right|^2\left[\frac{t}{2\pi}\frac{\sin^2\left(\frac{\omega_{fi}t}{2}\right)}{\left(\frac{\omega_{fi}t}{2}\right)^2}\right]. \qquad (П.20.7)$$

Так как в начальный момент времени $t=0$ $a_i(0)=1$, а $f_f^{(1)}=0$, то $\left|a_f^{(1)}(t)\right|^2$ можно рассматривать как вероятность перехода системы в течение времени t из состояния i в состояние f.

Покажем, что квадратная скобка в правой части этого выражения при $t\gg 1/\omega_{fi}$ ведет себя как δ-функция от ω_{fi}, т. е. для больших t

$$\left[\frac{t}{2\pi}\frac{\sin^2(\omega_{fi}t/2)}{(\omega_{fi}t/2)^2}\right]_{t\gg 1/\omega_{ii}}\longrightarrow\delta(\omega_{fi}). \qquad (П.20.8)$$

На рис. П.6 представлен множитель $\sin^2(\omega_{fi}t/2)/\omega_{fi}^2$ как функция ω_{fi}. При $\omega_{fi}=0$ квадратная скобка (П.20.8) равна $t/2\pi$, т. е. в единицах $1/\omega_{fi}$ очень велика. С другой стороны, полуширина центрального максимума, за пределами которого квадратная скобка практически равна нулю, очень мала, так как равна $2\pi/t$. Остается доказать, что интеграл от квадратной скобки по $d\omega_{fi}$ равен единице

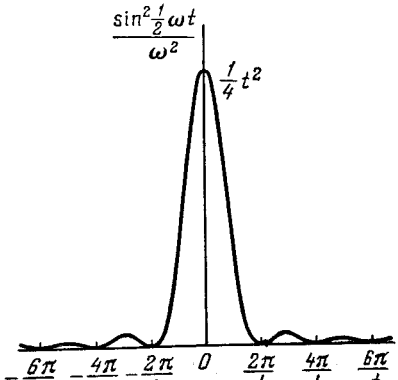


Рис. П. 6.

$$\int_{-\infty}^{\infty}\left[\frac{t}{2\pi}\frac{\sin^2(\omega_{fi}t/2)}{(\omega_{fi}t/2)^2}\right]d\omega_{fi}=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2 x}{x^2}\,dx=1,$$

так как последний интеграл равен π [2]).

---

[1]) Это утверждение неточно, так как предполагается, что $\mathscr{H}'=\text{const}$ для $t>0$, но $\mathscr{H}'=0$ для $t'<0$.

[2]) Бронштейн Н. Н., Семендяев К. А., с. 409.

Таким образом, вероятность перехода системы $i \longrightarrow f$ за единицу времени равна

$$w_{if} = \frac{|a_f^{(1)}(t)|^2}{t} = \frac{2\pi}{\hbar^2} |\mathscr{H}'_{fi}|^2 \delta(\omega_{fi}). \qquad (\text{П.20.9})$$

Так как [1])

$$\delta(\omega_{fi}) = \delta\left(\frac{\varepsilon_f - \varepsilon_i}{\hbar}\right) = \hbar\delta(\varepsilon_f - \varepsilon_i),$$

то

$$w_{if} = \frac{2\pi}{\hbar} |\mathscr{H}'_{fi}|^2 \delta(\varepsilon_f - \varepsilon_i). \qquad (\text{П.20.10})$$

Заметим, что структура квадратной скобки в (П.20.7), приводящейся к δ-функции, обусловливает пропорциональность $|a_f^{(1)}(t)|^2$ времени t, т. е. независимость w_{if} от времени.

Для определения полного числа переходов в единицу времени $\mathscr{W}$ надо просуммировать w_{if} по состояниям i и f

$$\mathscr{W} = \sum_{i,\,f} w_{if} = \frac{2\pi}{\hbar} \sum_{i,\,f} |\mathscr{H}'_{fi}|^2 \delta(\varepsilon_f - \varepsilon_i). \qquad (\text{П.20.10а})$$

3. Если возмущение $\mathscr{H}'(t)$ гармонически зависит от времени, т. е.

$$\mathscr{H}'(t) = \hat{\mathscr{H}}^0 (e^{-i\omega t} + e^{i\omega t}), \qquad (\text{П.20.11})$$

то из (П.20.5) следует

$$a_f^{(1)}(t) = \frac{1}{i\hbar} \int_0^t \mathscr{H}^0_{fi} [e^{i(\omega_{fi} - \omega)t'} + e^{i(\omega_{fi} + \omega)t'}]\, dt' =$$

$$= -\frac{\mathscr{H}^0_{fi}}{\hbar} \left[\frac{e^{i(\omega_{fi} - \omega)t} - 1}{\omega_{fi} - \omega} + \frac{e^{i(\omega_{fi} + \omega)t} - 1}{\omega_{fi} + \omega} \right]. \qquad (\text{П.20.12})$$

Это отличается от (П.20.6) тем, что имеем теперь в выражении для амплитуды $a_f^{(1)}(t)$ два слагаемых, причем первое отличается от (П.20.6) тем, что ω_{fi} заменено на $\omega_{fi} - \omega$, а второе — тем, что ω_{fi} заменено на $\omega_{fi} + \omega$. Первому слагаемому соответствует $\delta(\omega_{fi} - \omega)$, т. е. закон сохранения энергии $\varepsilon_f = \varepsilon_i + \hbar\omega$, при котором система *поглощает квант* $\hbar\omega$; второму слагаемому $-\delta(\omega_{fi} + \omega)$, т. е. закон сохранения энергии $\varepsilon_f = \varepsilon_i - \hbar\omega$, при котором система *испускает квант* $\hbar\omega$. Обе физические ситуации не могут выполняться одновременно, поэтому надо учитывать либо первое слагаемое в квадратной скобке (П.20.12), либо второе. Например, в случае поглощения кванта $\hbar\omega$ получим вместо (П.20.10)

$$w_{if} = \frac{2\pi}{\hbar} |\mathscr{H}^0_{fi}|^2 \delta(\varepsilon_f - \varepsilon_i - \hbar\omega). \qquad (\text{П.20.13})$$

В случае испускания кванта $\hbar\omega$ мы получим ту же формулу, но только перед $\hbar\omega$ будет стоять знак плюс.

Полное число квантов $\hbar\omega$ (фотонов), поглощенных в 1 *сек·в/см*^{3} кристалла, в результате перехода электронов из валентной зоны (v) в зону

[1]) Шифф Л., с. 67.

проводимости (c) равно

$$\mathscr{W}_{vc}=\frac{2\pi}{\hbar}\sum_{i,\,f}|\mathscr{H}^0_{fi}|^2\,\delta(\varepsilon_f-\varepsilon_i-\hbar\omega). \qquad (\text{П.20.13а})$$

Здесь суммирование ведется по заполненным состояниям (в 1 $см^3$) валентной зоны ($i\equiv v$) и по пустым состояниям зоны проводимости ($f\equiv c$), которые удовлетворяют закону сохранения энергии.

4. Рассмотрим второе приближение теории квантовых переходов (П.20.3а); это необходимо, например, в том случае, когда электрон переходит из валентной зоны в зону проводимости, поглощая при этом два кванта—фотон и фонон.

Пусть возмущение

$$\hat{\mathscr{H}}'(t)=\sum_k\hat{\mathscr{H}}^k e^{-i\omega_k t} \qquad (\text{П.20.14})$$

состоит из суммы членов, каждый из которых гармонически зависит от частоты ω_k. Мы рассматриваем случай, когда система поглощает кванты $\hbar\omega_k$, что не ограничивает общности рассмотрения.

Из (П.20.3а), (П.20.14) получим для перехода $i\longrightarrow$ через промежуточное состояние m

$$\frac{da_f^{(2)}}{dt}=\frac{1}{i\hbar}\sum_k\sum_m\mathscr{H}^k_{fm}a_m^{(1)}e^{i(\omega_{fm}-\omega_k)t}=$$

$$=-\frac{1}{i\hbar^2}\sum_{k,\,k'}\sum_m\mathscr{H}^k_{fm}\mathscr{H}^{k'}_{mi}e^{i(\omega_{fm}-\omega_k)t}\left\{\frac{e^{i(\omega_{mi}-\omega_{k'})t}-1}{\omega_{mi}-\omega_{k'}}\right\}=$$

$$=-\frac{1}{i\hbar^2}\sum_{k,\,k'}\sum_m\frac{\mathscr{H}^k_{fm}\mathscr{H}^{k'}_{mi}}{\omega_{mi}-\omega_{k'}}\left\{e^{i(\omega_{fi}-\omega_k-\omega_{k'})t}-e^{i(\omega_{fm}-\omega_k)t}\right\}. \qquad (\text{П.20.15})$$

Интегрируя, получим

$$a_f^{(2)}(t)=-\frac{1}{\hbar^2}\sum_{k,\,k'}\sum_m\frac{\mathscr{H}^k_{fm}\mathscr{H}^{k'}_{mi}}{\omega_{mi}-\omega_{k'}}\times$$

$$\times\left\{\frac{e^{i(\omega_{fi}-\omega_k-\omega_{k'})t}-1}{\omega_{fi}-\omega_k-\omega_{k'}}-\frac{e^{i(\omega_{fm}-\omega_k)t}-1}{\omega_{fm}-\omega_k}\right\}. \qquad (\text{П.20.16})$$

Это выражение состоит из суммы слагаемых, зависящих от m, k и k'. Из (П.20.6) видно, что только квадраты модулей отдельных слагаемых, входящих в фигурную скобку в (П.20.16), приводят к пропорциональности $|a_f^{(2)}(t)|^2$ от времени и к соответствующему закону сохранения энергии. Квадрат модуля первого слагаемого в фигурной скобке в (П.20.16) приводит к закону сохранения энергии: $\varepsilon_f=\varepsilon_i+\hbar\omega_k+\hbar\omega_{k'}$; в этих условиях можно пренебречь вторым «паразитным» слагаемым в фигурной скобке, которое появилось в результате нефизической ситуации—мгновенного включения возмущения в момент $t=0$ [1]. Мы приходим к заключению, что для вычисления $w_{if}=|a_f^{(2)}(t)|^2/t$ надо брать квадраты модулей отдельных слагаемых, входящих в правую часть (П.20.16) (отбросив второе слагаемое в фигурной скобке), так как только они обеспечивают пропорциональность $|a_f^{(2)}(t)|^2$ времени t и выполнение закона сохранения энергии между начальным и конечным состояниями.

---

[1]) Шифф Л., § 29.

Таким образом, мы для вероятности перехода $i \longrightarrow f$ через промежуточное состояние m получим

$$w_{if} = \frac{2\pi}{\hbar} \sum_{k,\,k'} \sum_{m} \frac{|\mathcal{H}^{k}_{fm}|^2 \, |\mathcal{H}^{k'}_{mi}|^2}{(\varepsilon_m - \varepsilon_i - \hbar\omega_{k'})^2} \, \delta(\varepsilon_f - \varepsilon_i - \hbar\omega_k - \hbar\omega_{k'}). \quad (\text{П.20.17})$$

Это выражение для определенного процесса приобретает более конкретный вид. В гл. VII, § 3 мы рассматриваем межзонные непрямые переходы. При этом возможен следующий двухступенчатый переход: валентный электрон вблизи $\boldsymbol{k} \approx 0$, вначале поглощает фотон, испытывая при этом прямой переход, а затем электрон поглощает фонон $\hbar\omega_q$, переходя в другую точку зоны Бриллюэна. Таким образом, переход $i \longrightarrow m$ происходит в результате взаимодействия с фотоном $\hbar\omega$, а это значит, что $\mathcal{H}^{k'}_{mi} = \mathcal{H}^{\text{фот}}_{mi}$; с другой стороны, переход $m \longrightarrow f$ связан с взаимодействием с фононом $\hbar\omega_q$, поэтому $\mathcal{H}^{k}_{fm} = \mathcal{H}^{\text{фон}}_{fm}$.

В результате для этого перехода (П.20.17) приобретает вид

$$w_{if} = \frac{2\pi}{\hbar} \sum_{m} \frac{|\mathcal{H}^{\text{фон}}_{fm}|^2 \, |\mathcal{H}^{\text{фот}}_{mi}|^2}{(\varepsilon_m - \varepsilon_i - \hbar\omega)^2} \, \delta(\varepsilon_f - \varepsilon_i - \hbar\omega_q - \hbar\omega). \quad ((\text{П.20.18})$$

Выражение для полного числа переходов $\mathscr{W}$ получается из (П.20.18), аналогично (П.20.13а) суммированием по начальному состоянию i и по конечному состоянию f.

Приложение 21

Запишем подробнее, используя (3.1) и (3.3), матричный элемент (3.7а):

$$\langle \boldsymbol{k}',\, N'_{qj} \,|\, \Delta V \,|\, \boldsymbol{k},\, N_{qj} \rangle =$$

$$= -\frac{1}{\sqrt{NM}} \int \left[\psi^*_{\boldsymbol{k}'}(\boldsymbol{r}) \prod_{qj} \psi^*_{N'_{qj}}(Q_{qj}) \operatorname{grad} V \sum_{qj}{}' e_{qj} \left(a_{qj} e^{iqr} + a^*_{qj} e^{-iqr} \right) \times \right.$$

$$\left. \times \psi_{\boldsymbol{k}}(\boldsymbol{r}) \prod_{qj} \psi_{Nqj}(Q_{qj}) \right] d\tau \prod_{qj} dQ_{qj}. \quad (\text{П.21.1})$$

Изменим порядок интегрирования и суммирования и выделим интегралы по электронным координатам и нормальным колебаниям решетки:

$$-\frac{1}{\sqrt{NM}} \sum_{qj}{}' \left\{ \left[e_{qj} \frac{1}{N} \int \operatorname{grad} V e^{i(\boldsymbol{k}+\boldsymbol{q}-\boldsymbol{k}')\boldsymbol{r}} u_{\boldsymbol{k}}(\boldsymbol{r}) u^*_{\boldsymbol{k}'}(\boldsymbol{r}) \, d\tau \right] \times \right.$$

$$\times \left[\int \prod_{qj} \psi^*_{N'_{qj}}(Q_{qj}) a_{qj} \prod_{qj} \psi_{Nqj}(Q_{qj}) (dQ_{qj}) \right] +$$

$$+ \left[e_{qj} \frac{1}{N} \int \operatorname{grad} V e^{i(\boldsymbol{k}-\boldsymbol{q}-\boldsymbol{k}')\boldsymbol{r}} u_{\boldsymbol{k}}(\boldsymbol{r}) u^*_{\boldsymbol{k}'}(\boldsymbol{r}) \, d\tau \right] \times$$

$$\left. \times \left[\int \prod_{qj} \psi^*_{N'_{qj}}(Q_{qj}) a^*_{qj} \prod_{qj} \psi_{Nqj}(Q_{qj}) (dQ_{qj}) \right] \right\}, \quad (\text{П.21.2})$$

где $(dQ_{qj}) \equiv \prod_{qj} dQ_{qj}$.

1-ю, 2-ю, 3-ю и 4-ю квадратные скобки в (П.21.2) обозначим последовательно K^+, L, K^- и L^*. Рассмотрим вначале L. Интегралы от всех пар

$\Psi^*_{N'_{qj}}(Q_{qj})\,\Psi_{N_{qj}}(Q_{qj})$ для Q_{qj}, не соответствующих a_{qj}, равны единице, если $N'_{qj}=N_{qj}$, и нулю в остальных случаях. Матричный элемент от a_{qj}, согласно (III.10.25), отличен от нуля только при $N'_{qj}=N_{qj}-1$, и в этом случае

$$L=\langle N_{qj}-1\,|\,a_{qj}\,|\,N_{qj}\rangle=\sqrt{\frac{\hbar N_{qj}}{2\omega_{qj}}}. \qquad (\text{П.21.3})$$

При этом $L^*=0$. Рассуждая точно так же, убедимся, что L^* отлично от нуля только при $N'_{qj}=N_{qj}+1$, и в этом случае (III.10.26)

$$L^*=\langle N_{qj}+1\,|\,a^*_{qj}\,|\,N_{qj}\rangle=\sqrt{\frac{\hbar(N_{qj}+1)}{2\omega_{qj}}}. \qquad (\text{П.21.3а})$$

Таким образом, при взаимодействии электронов с колебаниями осуществляются только такие переходы, при которых число фононов одного сорта qj или уменьшается или увеличивается на единицу, а число всех остальных фононов остается без изменения.

Рассмотрим теперь интеграл K^+ по электронным координатам r. Заменим в K^+ r на a_n+r', где r' меняется в пределах одной элементарной ячейки. Учитывая, что $V(r)$, $u_k(r)$ и $u^*_{k'}(r)$ трехмерно периодичны с периодами решетки, получим

$$K^+=\frac{1}{N}\sum_n e^{i(k+q-k')a_n}\int e^{i(k+q-k')r}\,e_{qj}\,\operatorname{grad}V(r)\,u_k(r)\,u^*_{k'}(r)\,d\tau_0, \qquad (\text{П.21.4})$$

где интеграл берется по элементарной ячейке и мы опустили штрих у переменной r' под знаком интеграла. Согласно (П.6.4) сумма по n отлична от нуля и равна N, если

$$k'=k+q. \qquad (\text{П.21.4а})$$

Однако в данном случае надо учесть следующую возможность. Сумма векторов $k+q$ может не соответствовать точке в первой бриллюэновской зоне. В этом случае для приведения вектора k' к 1-й бриллюэновской зоне надо положить

$$k'=k+q+b_g, \qquad (\text{П.21.4б})$$

где b_g — вектор обратной решетки. При этом сумма по n в (П.21.4) тоже отлична от нуля.

Рассеяние, соответствующее (П.21.4б), называется *процессом переброса* (umklapp-prozess), оно играет существенную роль в установлении теплового равновесия в решетке (Пайерлс).

В кинетических явлениях, рассматриваемых далее, процессы переброса роли не играют, поэтому в дальнейшем мы будем их игнорировать.

В случае (П.21.4а)

$$K^+=\int e_{qj}\,\operatorname{grad}V u_k(r)\,u^*_{k'}(r)\,d\tau_0=\int u_k u^*_{k'}\frac{\partial V}{\partial s}\,d\tau_0=$$
$$=\int\frac{\partial}{\partial s}\left(u_k u^*_{k'}V\right)d\tau_0-\int V\frac{\partial}{\partial s}\left(u_k u^*_{k'}\right)d\tau_0, \qquad (\text{П.21.5})$$

где $\boldsymbol{k}'=\boldsymbol{k}+\boldsymbol{q}$, а $\partial/\partial s$ означает дифференцирование по направлению $\boldsymbol{e}_{qj}$. Первый интеграл правой части (П.21.5) может быть преобразован к интегралу по поверхности [1])

$$\int \frac{\partial}{\partial s}\left(u_k u^*_{k'} V\right) d\tau_0 = \boldsymbol{e}_{qj} \int \operatorname{grad}\left(u_k u^*_{k'} V\right) d\tau_0 = \boldsymbol{e}_{qj} \oint \boldsymbol{\nu} u_k u^*_{k'} V \, d\sigma_0, \quad \text{(П.21.5а)}$$

где $\boldsymbol{\nu}$ — внешняя нормаль к поверхности элементарной ячейки. Так как в соответствующих точках противоположных граней элементарной ячейки $u_k u^*_{k'} V$ одинаково, а $\boldsymbol{e}_{qj}\boldsymbol{\nu} = \cos(\boldsymbol{\nu}, \boldsymbol{e}_{qj})$ одинаково по величине, но противоположно по знаку, то весь интеграл (П.21.5а) равен нулю.

Таким образом,

$$K^+ = -\int V \frac{\partial}{\partial s}\left(u_k \, u^*_{k'}\right) d\tau_0. \quad \text{(П.21.5б)}$$

Функции u_k и $u^*_{k'}$ удовлетворяют уравнению (IV.3.8)

$$-\frac{\hbar^2}{2m}\nabla^2 u_k + V(\boldsymbol{r})\, u_k - \frac{i\hbar^2}{m}(\boldsymbol{k}\operatorname{grad} u_k) = \left(\varepsilon_k - \frac{\hbar^2 k^2}{2m}\right) u_k,$$
$$-\frac{\hbar^2}{2m}\nabla^2 u^*_{k'} + V(\boldsymbol{r})\, u^*_{k'} + \frac{i\hbar^2}{m}(\boldsymbol{k}' \operatorname{grad} u^*_{k'}) = \left(\varepsilon_{k'} - \frac{\hbar^2 k'^2}{2m}\right) u^*_{k'}.$$

Умножая первое уравнение на $\partial u^*_{k'}/\partial s$, второе — на $\partial u_k/\partial s$, складывая и интегрируя по элементарной ячейке, получим

$$-\int V \frac{\partial}{\partial s}\left(u_k \, u^*_{k'}\right) d\tau_0 = \int \left\{ -\frac{\hbar^2}{2m}\left[\frac{\partial u^*_{k'}}{\partial s}\nabla^2 u_k + \frac{\partial u_k}{\partial s}\nabla^2 u^*_{k'}\right] - \right.$$
$$-\frac{i\hbar^2}{m}\left[\frac{\partial u^*_{k'}}{\partial s}(\boldsymbol{k}\operatorname{grad} u_k) - \frac{\partial u_k}{\partial s}(\boldsymbol{k}' \operatorname{grad} u^*_{k'})\right] -$$
$$\left. -\left[\left(\varepsilon_k - \frac{\hbar^2 k^2}{2m}\right) u_k \frac{\partial u^*_{k'}}{\partial s} + \left(\varepsilon_{k'} - \frac{\hbar^2 k'^2}{2m}\right) u^*_{k'} \frac{\partial u_k}{\partial s}\right]\right\} d\tau_0. \quad \text{(П.21.6)}$$

Преобразуем вторые слагаемые в каждой из квадратных скобок, используя самосопряженность операторов ∇^2 и $i\nabla$ и равенство нулю интегралов вида (П.21.5а)

$$1)\ \int \frac{\partial u_k}{\partial s}\nabla^2 u^*_{k'} d\tau_0 = \int u^*_{k'}\nabla^2\left(\frac{\partial u_k}{\partial s}\right) d\tau_0 = \int u^*_{k'} \frac{\partial}{\partial s}(\nabla^2 u_k)\, d\tau_0 =$$
$$= \int \frac{\partial}{\partial s}\left(u^*_{k'}\nabla^2 u_k\right) d\tau_0 - \int \frac{\partial u^*_{k'}}{\partial s}\nabla^2 u_k \, d\tau_0 = -\int \frac{\partial u^*_{k'}}{\partial s}\nabla^2 u_k \, d\tau_0;$$

$$2)\ -\int \frac{\partial u_k}{\partial s} i\nabla u^*_{k'} d\tau_0 = \int u^*_{k'}(+i)\nabla \frac{\partial u_k}{\partial s} d\tau_0 = \int u^*_{k'} \frac{\partial}{\partial s}(+i\nabla u_k)\, d\tau_0 =$$
$$= \int \frac{\partial}{\partial s}\left[u^*_{k'}(+i)\nabla u_k\right] d\tau_0 - \int \frac{\partial u^*_{k'}}{\partial s} i\nabla u_k \, d\tau_0 = -\int \frac{\partial u^*_{k'}}{\partial s} i\nabla u_k \, d\tau_0;$$

$$3)\ \int u^*_{k'} \frac{\partial u_k}{\partial s} d\tau_0 = \int \frac{\partial}{\partial s}\left(u^*_{k'} u_k\right) d\tau_0 - \int u_k \frac{\partial u^*_{k'}}{\partial s} d\tau_0 = -\int u_k \frac{\partial u^*_{k'}}{\partial s} d\tau_0.$$

Используя эти три равенства, получим из (П.21.5б) и (П.21.6)

$$K^+ = -\frac{\hbar^2 i}{m}\left[(\boldsymbol{k}-\boldsymbol{k}')\int \operatorname{grad} u_{\boldsymbol{k}}\frac{\partial u^*_{\boldsymbol{k}'}}{\partial s}\,d\tau_0\right] - \\ -\left[\varepsilon_{\boldsymbol{k}}-\varepsilon_{\boldsymbol{k}'}-\left(\frac{\hbar^2 k^2}{2m}-\frac{\hbar^2 k'^2}{2m}\right)\right]\int u_{\boldsymbol{k}}\frac{\partial u^*_{\boldsymbol{k}'}}{\partial s}\,d\tau_0. \qquad (\text{П.21.7})$$

Если эффективная масса m^* равна массе электрона m, то квадратная скобка при втором слагаемом равна нулю, так как $\varepsilon_{\boldsymbol{k}} = \hbar^2 k^2/2m$ и $\varepsilon_{\boldsymbol{k}'} = \hbar^2 k'^2/2m$. Если эффективная масса порядка m, то эта квадратная скобка порядка $\hbar^2 k^2/m$. Интеграл в квадратной скобке первого слагаемого порядка

$$\int \operatorname{grad} u_{\boldsymbol{k}}\frac{\partial u^*_{\boldsymbol{k}}}{\partial s}\,d\tau_0 \approx \frac{1}{a}\int u_{\boldsymbol{k}}\frac{\partial u^*_{\boldsymbol{k}'}}{\partial s}\,d\tau_0,$$

где a — постоянная решетки, поэтому отношение второго слагаемого к первому порядка

$$\frac{\hbar^2 k^2}{m} : \frac{\hbar^2 k}{ma} = ak,$$

что для полупроводников много меньше единицы.

Эта оценка имеет место, если экстремум энергии электрона находится в центре бриллюэновской зоны в $\boldsymbol{k}=0$.

Пренебрегая вторым слагаемым в (П.21.7) и используя (П.21.4а), получим

$$K^+ = \frac{\hbar^2 i}{m}\boldsymbol{q}\int \operatorname{grad} u_{\boldsymbol{k}}\int \frac{\partial u^*_{\boldsymbol{k}'}}{\partial s}\,d\tau_0 = \frac{\hbar^2 i}{m}\sum_{\alpha,\,\beta} q_\alpha e_\beta \int \frac{\partial u_{\boldsymbol{k}}}{\partial x_\alpha}\frac{\partial u^*_{\boldsymbol{k}'}}{\partial x_\beta}\,d\tau_0, \qquad (\text{П.21.8})$$

где e_β — прямоугольная составляющая вектора поляризации $\boldsymbol{e}_{qj}$. Здесь использовано соотношение

$$\frac{\partial u^*_{\boldsymbol{k}'}}{\partial s} = \boldsymbol{e}_{qj}\operatorname{grad} u^*_{\boldsymbol{k}'}.$$

В кубическом кристалле $u_{\boldsymbol{k}}$ — либо четная, либо нечетная функция x_α, поэтому $\partial u_{\boldsymbol{k}}/\partial x_\alpha$ — в первом случае нечетная функция, а во втором — четная от x_α. При интегрировании в (П.21.8) по x_α от $-a/2$ до $+a/2$ (a — постоянная решетки) интеграл для $\beta \neq \alpha$ равен нулю, поэтому

$$K^+ = \frac{\hbar^2 i}{m}\sum_{\alpha} q_\alpha e_\alpha \int \frac{\partial u_{\boldsymbol{k}}}{\partial x_\alpha}\frac{\partial u^*_{\boldsymbol{k}'}}{\partial x_\alpha}\,d\tau_0.$$

Можно показать, что $u_{\boldsymbol{k}}$ слабо зависит от $\boldsymbol{k}$, так что $u_{\boldsymbol{k}} \approx u_{\boldsymbol{k}'} \equiv u$. Так как все три прямоугольные оси в кубическом кристалле эквивалентны, то

$$\int \frac{\partial u_{\boldsymbol{k}}}{\partial x_\alpha}\frac{\partial u_{\boldsymbol{k}'}}{\partial x_\alpha}\,d\tau_0 \approx \int \left|\frac{\partial u}{\partial x_\alpha}\right|^2 d\tau_0 = \frac{1}{3}\int |\operatorname{grad} u|^2\,d\tau_0$$

для всех α.

Таким образом,

$$K^+ = \frac{\hbar^2 i}{m}\frac{1}{3}\int |\operatorname{grad} u|^2\,d\tau_0\,(\boldsymbol{q}\boldsymbol{e}_{qj}). \qquad (\text{П.21.8а})$$

---

[1]) Смирнов В. И. т. 2, § 11.

Рассматривая три ветви колебаний: одну продольную ($j=1$) $\boldsymbol{e}_{q1} \parallel \boldsymbol{q}$ и две поперечные ($j=2, 3$) $\boldsymbol{e}_{q2}$ и $\boldsymbol{e}_{q3} \perp \boldsymbol{q}$, видим, что при сделанных нами приближениях с электроном взаимодействуют только продольные колебания.

В результате (П.21.8а) равно

$$K^+ = i\frac{2}{3} qC, \tag{П.21.8б}$$

где

$$C = \frac{\hbar^2}{2m}\int |\operatorname{grad} u|^2 d\tau_0. \tag{П.21.8в}$$

Объединяя (П.21.3) и (П.21.8б), получим выражение (3.10) текста, равное матричному элементу, связанному с поглощением фонона.

Рассматривая случай, когда $L^* \neq 0$ и равно (П.21.3а), т. е. испускание электроном фонона, легко видеть, что все отличие при вычислении K^- по сравнению с K^+ заключается в том, что в этом случае

$$\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{q} \tag{П.21.9}$$

вместо (П.21.4а). В результате K^- отличается от K^+ только знаком. Объединяя (П.21.3а) с выражением для K^-, получим (3.10а) текста.

Приложение 22

Докажем соотношение (5.15), рассчитав непосредственно из квантовой механики изменение энергии электрона у дна зоны проводимости при деформации кристалла [1]).

При однородном всестороннем растяжении кубического кристалла постоянная решетки

$$a = a_0(1+\varepsilon), \tag{П.22.1}$$

где a_0 — постоянная решетки недеформированного кристалла и $\varepsilon = \varepsilon_{ii} = \dfrac{\partial u_i}{\partial x_i}$.

Если $\Delta\mathscr{E}$ — понижение нижнего края зоны проводимости, связанное с этим растяжением, то по определению (5.4)

$$\mathscr{E}_1 = -\frac{\Delta\mathscr{E}}{3\varepsilon}. \tag{П.22.2}$$

В однородном деформированном кристалле, так же как и в недеформированном, электрон проводимости описывается блоховской функцией

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = u_{\boldsymbol{k}}(\boldsymbol{r})\, e^{i\boldsymbol{k}\boldsymbol{r}}, \tag{П.22.3}$$

где $u_{\boldsymbol{k}}(\boldsymbol{r})$ — трехмерно периодическая функция, обладающая периодом a. У дна зоны проводимости при $\boldsymbol{k}=0$ модулирующая функция $u_{\boldsymbol{k}=0} \equiv u_0$ удовлетворяет уравнению (IV.3.8)

$$-\frac{\hbar^2}{2m}\nabla^2 u_0 + V(\boldsymbol{r})\, u_0 = \mathscr{E} u_0, \tag{П.22.4}$$

где $V(\boldsymbol{r})$ — периодический потенциал, действующий на электрон, а $\mathscr{E}$ — энергия нижнего края зоны проводимости в деформированном кристалле.

Для применения теории возмущений необходимо, чтобы возмущенная и невозмущенная волновые функции удовлетворяли одним и тем же граничным условиям, которые для функции u_0 заменяются условиями периодичности. Для того чтобы u_0 в деформированном и недеформированном кристаллах удовлетворяла одним и тем же условиям периодичности, введем безразмерные

---

[1]) Пикус Г. Е. — ЖТФ, 1958, т. 28, с 2390.

координаты

$$x' = \frac{x}{a} = \frac{x}{a_0(1+\varepsilon)} \qquad (П.22.5)$$

и аналогично для y' и z'. Очевидно, что при увеличении x на $a = a_0(1+\varepsilon)$ как в деформированном ($\varepsilon \neq 0$), так и в недеформированном ($\varepsilon = 0$) кристаллах координата x' меняется на 1, поэтому период u_0 в безразмерных координатах x' в обоих случаях один и тот же и равен 1. Переходя к безразмерным координатам $\boldsymbol{r}' = \boldsymbol{r}/a$, получим вместо (П.22.4)

$$-\frac{\hbar^2}{2m} a^{-2} \nabla^2_{\boldsymbol{r}'} u_0 + V(a\boldsymbol{r}')\, u_0 = \mathscr{E} u_0. \qquad (П.22.6)$$

Обозначая через V_0 и $\mathscr{E}_0$ периодический потенциал и энергию электрона (при $\boldsymbol{k}=0$) в недеформированном кристалле, получим аналогично (П.22.6)

$$-\frac{\hbar^2}{2m} a_0^{-2} \nabla^2_{\boldsymbol{r}'} u_0 + V_0(a_0\boldsymbol{r}')\, u_0 = \mathscr{E}_0 u_0, \qquad (П.22.7)$$

где $\boldsymbol{r}' = \boldsymbol{r}/a_0$.

Во всех реальных случаях деформация $\varepsilon \ll 1$, поэтому вычислим смещения нижнего края зоны проводимости по теории возмущений в первом порядке по ε, т. е. пренебрегая поправками порядка ε^2, и т. д.

Из теории возмущений квантовой механики известно, что для того, чтобы определить энергию возмущения в первом приближении, волновые функции достаточно взять в нулевом приближении. Это означает, что u_0 в (П.22.6) и (П.22.7) можно считать одинаковыми. Умножая (П.22.6) и (П.22.7) слева на $u_0^*(\boldsymbol{r}/a_0)$, вычитая полученные равенства почленно и интегрируя по $d\tau' = d\tau/a_0^3$, получим

$$(\mathscr{E}-\mathscr{E}_0)\int u_0^* u_0\, d\tau' = \int u_0^* \left\{ -\frac{\hbar^2}{2m} a_0^{-2}(1+\varepsilon)^{-2} \nabla^2_{\boldsymbol{r}'} + V[a_0(1+\varepsilon)\boldsymbol{r}'] + \right.$$
$$\left. + \frac{\hbar^2}{2m} a_0^{-2} \nabla^2_{\boldsymbol{r}'} - V_0(a_0\boldsymbol{r}') \right\} u_0\, d\tau'. \qquad (П.22.8)$$

Если нормировать блоховскую функцию (П.22.3) согласно (IV.3.6), то

$$\int u_0^* u_0\, d\tau' = \frac{1}{a_0^3}\int |u_0|^2\, d\tau = \frac{1}{a_0^3}. \qquad (П.22.9)$$

Разложим первое слагаемое в фигурной скобке (П.22.8) в ряд по ε и ограничимся членами нулевого и первого порядка; сократим член нулевого порядка с третьим слагаемым в фигурных скобках; перейдем обратно к переменной $\boldsymbol{r} = \boldsymbol{r}' a_0$; используя тогда (П.22.9), получим из (П.22.8)

$$\mathscr{E}-\mathscr{E}_0 = \Delta\mathscr{E} = \varepsilon\frac{\hbar^2}{m}\int u_0^* \nabla^2 u_0\, d\tau + \int u_0^* \{V[(1+\varepsilon)\boldsymbol{r}] - V_0(\boldsymbol{r})\}\, u_0\, d\tau. \qquad (П.22.10)$$

Применяя формулу Грина [1]), получим

$$\int u_0^* \nabla^2 u_0\, d\tau = -\int |\nabla u_0|^2\, d\tau + \oint u_0^* (\nabla u_0\, d\sigma), \qquad (П.22.11)$$

где последний интеграл, взятый по поверхности элементарной ячейки, равен нулю.

Согласно гипотезе деформируемых ионов (3.3) $V[(1+\varepsilon)\boldsymbol{r}] = V_0(\boldsymbol{r})$; тогда из (П.22.10), (П.22.11) и (П.22.2), если воспользоваться определением C

[1]) Смирнов В. И., т. 2, гл. VII, п. 203.

(3.10б), следует

$$\mathscr{E}_1 = -\frac{\Delta\mathscr{E}}{3\varepsilon} = \frac{\hbar^2}{3m}\int |\nabla u_0|^2\, d\tau = \frac{2}{3}C, \qquad (\text{П.22.12})$$

что совпадает с (5.15).

Приложение 23

Определение χ_n (или χ_p) из уравнения (2.8) (или (2.10)) сводится к решению векторного уравнения

$$\boldsymbol{x} = \boldsymbol{a} + [\boldsymbol{bx}] \qquad (\text{П.23.1})$$

относительно неизвестного вектора $\boldsymbol{x}$.

Заметим, что из (П.23.1) следует

$$\boldsymbol{bx} = \boldsymbol{ba},$$

так как

$$(\boldsymbol{b}\,[\boldsymbol{bx}]) = 0.$$

Подставим вместо $\boldsymbol{x}$ в правую часть (П.23.1) его выражение $\boldsymbol{a} + [\boldsymbol{bx}]$, тогда

$$\boldsymbol{x} = \boldsymbol{a} + [\boldsymbol{ba}] + [\boldsymbol{b}\,[\boldsymbol{bx}]]. \qquad (\text{П.23.2})$$

Используя тождественное преобразование [1])

$$[\boldsymbol{b}\,[\boldsymbol{bx}]] = \boldsymbol{b}(\boldsymbol{bx}) - \boldsymbol{x}b^2$$

и (П.23.2), получим

$$\boldsymbol{x} = \boldsymbol{a} + [\boldsymbol{ba}] + \boldsymbol{b}(\boldsymbol{ba}) - \boldsymbol{x}b^2,$$

откуда

$$\boldsymbol{x} = \frac{\boldsymbol{a} + [\boldsymbol{ba}] + (\boldsymbol{ab})\,\boldsymbol{b}}{1 + b^2}. \qquad (\text{П.23.3})$$

Применяя это выражение, получим (2.11) и (2.12).

Приложение 24

Индексы μ, α и ν в символе $\delta_{\mu\alpha\nu}$ независимо принимают значения 1, 2 и 3. Рассмотрим все возможные *перестановки* чисел 1, 2 и 3:

$$(123),\quad (231),\quad (312),\quad (132),\quad (213),\quad (321). \qquad (\text{П.24.1})$$

Их число равно $3! = 6$.

Назовем *беспорядком* в перестановке тот факт, что бо́льшее число стоит впереди меньшего, и подсчитаем число беспорядков в перестановках (П.24.1). Легко видеть, что оно равно

$$0,\quad 2,\quad 2,\quad 1,\quad 1,\quad 3, \qquad (\text{П.24.2})$$

т. е. для первых трех перестановок число беспорядков четное, для трех последних — нечетное.

Если мы положим:

$\delta_{\mu\alpha\nu} = +1$, когда $(\mu\alpha\nu)$ образует четное число беспорядков,

$\delta_{\mu\alpha\nu} = -1$, когда $(\mu\alpha\nu)$ образует нечетное число беспорядков,

$\delta_{\mu\alpha\nu} = 0$, когда среди $(\mu\alpha\nu)$ имеются одинаковые индексы, то можно непосредственно убедиться в том, что векторное произведение $[\boldsymbol{vH}]$ может быть посредством символа $\delta_{\mu\alpha\nu}$ записано в виде (8.22).

---

[1]) Смирнов В. И., т. III, ч. 1.

Назовем *транспозицией* операцию, при которой в некоторой перестановке обмениваются местами два элемента (индекса). Легко убедиться в том, что транспозиция меняет число беспорядков на нечетное число. Таким образом, перестановки с четным числом беспорядков в результате транспозиции переходят в перестановки с нечетным числом беспорядков и наоборот, так что

$$\delta_{\mu\alpha\nu} = -\delta_{\alpha\mu\nu} \qquad (\text{П.24.3})$$

и т. д.

Циклическая перестановка индексов $\mu \longrightarrow \alpha$, $\alpha \longrightarrow \nu$, $\nu \longrightarrow \mu$ символ $\delta_{\mu\alpha\nu}$ не меняет.

Очевидно, что $\delta^2_{\mu\alpha\nu} = 1$ или 0.

Можно показать, что $\delta_{\mu\alpha\nu}$—тензор 3-го ранга, у которого из 27 компонент отличны от нуля только шесть.

Изложенные выше понятия и результаты могут быть легко обобщены на случай перестановок из n элементов [1]).

---

[1]) Смирнов В. И., т. II, п. 117.

ОГЛАВЛЕНИЕ

Из предисловия к первому изданию 6
Предисловие ко второму изданию 7

Глава I. **Геометрия кристаллических решеток и дифракция рентгеновских лучей** 9
§ 1. Простые и сложные кристаллические решетки 9
§ 2. Примеры конкретных кристаллических структур 15
§ 3. Прямая и обратная решетки кристалла 20
§ 4. Формулы Лауэ и Вульфа—Брэгга для дифракции рентгеновских лучей в кристалле. Атомный и структурный факторы рассеяния 24

Глава II. **Элементы теории групп и симметрия кристаллов** 30
§ 1. Введение 30
§ 2. Элементы абстрактной теории групп 33
§ 3. Точечные группы 39
§ 4. Группа трансляций. Сингонии (кристаллические системы) и решетки Браве 48
§ 5. Кристаллические классы. Пространственные группы 55
§ 6. Неприводимые представления групп и теория характеров . . 64
§ 7. Квантовая механика и теория групп 79
§ 8. Применение теории групп к исследованию расщепления уровней энергии примесного атома в кристалле и к классификации нормальных колебаний многоатомной молекулы 86
§ 9. Применение теории групп к трансляционной симметрии кристалла 96
§ 10. Правила отбора 106

Глава III. **Колебания атомов кристаллической решетки** 110
§ 1. Природа сил взаимодействия атомов в кристалле 110
§ 2. Колебания и волны в простой одномерной (линейной) решетке 119
§ 3. Колебания и волны в сложной одномерной (линейной) решетке 125
§ 4. Нормальные координаты для простой одномерной решетки . 130
§ 5. Колебания атомов трехмерной сложной кристаллической решетки 133
§ 6. Нормальные координаты колебаний кристаллической решетки 145
§ 7. Колебания простой кубической решетки 151
§ 8. Применение теории групп к исследованию нормальных колебаний кристаллической решетки 157
§ 9. Колебания и волны в кристаллах в приближении изотропного континуума 166
§ 10. Квантование колебаний кристаллической решетки. Фононы . 173
§ 11. Теория теплоемкости кристаллической решетки 178
§ 12. Уравнение состояния твердого тела 186
§ 13. Тепловое расширение и теплопроводность твердого тела . . 191

Глава IV. **Электроны в идеальном кристалле** 196
§ 1. Общая постановка задачи. Адиабатическое приближение . . 196
§ 2. Метод Хартри—Фока 199
§ 3. Электрон в периодическом поле 206
§ 4. Понятие о положительных дырках почти заполненной валентной зоны 217
§ 5. Приближение почти свободных (слабо связанных) электронов 221
§ 6. Зоны Бриллюэна 225
§ 7. Приближение сильно связанных электронов 231
§ 8. Структура энергетических зон и симметрия волновых функций в простой кубической решетке и в кристалле сурьмянистого индия 246
§ 9. Группы волнового вектора для решетки типа германия . . . 253
§ 10. Спин-орбитальное взаимодействие и двойные группы . . . 258
§ 11. Двойные группы в кристаллах InSb и Ge 266
§ 12. Спин-орбитальное расщепление в кристаллах InSb и Ge . . 272
§ 13. Исследование спектра электронов (дырок) вблизи минимума (максимума) энергии в зоне Бриллюэна (***kp***-метод) 276
§ 14. Симметрия, связанная с обращением времени 291
§ 15. Структура энергетических зон некоторых полупроводников . 299

Глава V. **Локализованные состояния электрона в кристалле** 304
§ 1. Функции Ванье. Движение электрона в поле примеси 304
§ 2. Локализованные состояния электрона в неидеальной решетке 311
§ 3. Экситоны 318
§ 4. Поляроны 325

Глава VI. **Электрические, тепловые и магнитные свойства твердых тел** 336
§ 1. Металлы, диэлектрики и полупроводники 336
§ 2. Статистическое равновесие свободных электронов в полупроводниках и металлах 338
§ 3. Теплоемкость свободных электронов в металлах и полупроводниках 349
§ 4. Магнитные свойства вещества. Парамагнетизм газов и электронов проводимости в металлах и полупроводниках 352
§ 5. Диамагнетизм атомов и электронов проводимости. Магнитные свойства полупроводников 361
§ 6. Циклотронный (диамагнитный) резонанс 372
§ 7. Контакт полупроводника с металлом. Выпрямление 380
§ 8. Свойства p—n-переходов 387
§ 9. Генерация и рекомбинация носителей тока. Квазиуровни Ферми 394

Глава VII. **Оптика полупроводников** 398
§ 1. Дисперсионные соотношения Крамерса—Кронига 398
§ 2. Межзонное поглощение света, связанное с прямыми переходами 403
§ 3. Межзонные непрямые переходы 417
§ 4. Поглощение света в полупроводниках свободными носителями 426
§ 5. Поляритоны 428
§ 6. Эффект вращения Фарадея 432
§ 7. Теория межзонного поглощения света в квантующем магнитном поле 436
§ 8. Поглощение света в полупроводниках в однородном электрическом поле (эффект Франца—Келдыша) 445

Глава VIII. Кинетическое уравнение и время релаксации для электронов проводимости в кристаллах 454
§ 1. Явления переноса и кинетическое уравнение Больцмана . . . 454
§ 2. Кинетическое уравнение для электронов в кристалле 463
§ 3. Рассеяние электронов на акустических колебаниях решетки . 467
§ 4. Время релаксации электронов проводимости в атомном полупроводнике и металле 471
§ 5. Теория деформационного потенциала в кубических кристаллах с простой зонной структурой 476
§ 6. Рассеяние электронов проводимости в ионных кристаллах на колебаниях решетки 481
§ 7. Рассеяние электронов проводимости на заряженных и нейтральных атомах примесей 488

Глава IX. Кинетические процессы (явления переноса) в полупроводниках 494
§ 1. Введение 494
§ 2. Определение неравновесной функции для электронов проводимости в случае сферически-симметричной зоны 497
§ 3. Электропроводность невырожденных полупроводников с простой зонной структурой 502
§ 4. Термоэлектрические явления в невырожденных полупроводниках с простой зонной структурой 506
§ 5. Гальваномагнитные явления в невырожденных полупроводниках с простой зонной структурой 513
§ 6. Термомагнитные явления в невырожденных полупроводниках с простой зонной структурой 520
§ 7. Явления переноса в полупроводниках с простой зоной при произвольном вырождении 527
§ 8. Явления переноса в полупроводниках типа германия и кремния 533
§ 9. Явления переноса в полупроводниках со сферической непараболической зоной 553
§ 10. Эффект «фононного увлечения» в полупроводниках 557
§ 11. Квантовая теория гальвано- и термомагнитных явлений в полупроводниках 565

Приложения 574

Андрей Иванович Ансельм

ВВЕДЕНИЕ В ТЕОРИЮ ПОЛУПРОВОДНИКОВ

М., 1978 г., 616 стр. с илл.