

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ имени
М.В.ЛОМОНОSOBA**

**Научно-исследовательский институт
ядерной физики имени Д.В.Скобельцына**

**А. Н. Васильев
В. В. Михайлин**

**Введение в
спектроскопию
диэлектриков**

Часть I

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ имени М.В.ЛОМОНОСОВА

НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ
ИНСТИТУТ ЯДЕРНОЙ ФИЗИКИ
имени Д.В.СКОБЕЛЬЦЫНА

А.Н.Васильев, В.В.Михайлин

**Введение в
спектроскопию диэлектриков**

Часть I

Москва

2008

УДК 535, 539.1
ББК 22.344

А.Н.Васильев, В.В.Михайлин
Введение в спектроскопию диэлектриков, часть I
(учебное пособие)

В учебном пособии отражены как теоретические, так и экспериментальные аспекты спектроскопии твердого тела. В первой части пособия рассматриваются как общие подходы к описанию взаимодействия оптического излучения с конденсированными средами, так и элементарные электронные и фоновые возбуждения. Пособие написано на основе курсов лекций, читаемых авторами студентам старших курсов на физическом факультете Московского государственного университета имени М.В.Ломоносова.

Учебное пособие рекомендовано к печати учебно-методической комиссией кафедры оптики и спектроскопии физического факультета МГУ.

© МГУ
© НИИЯФ МГУ
© А.Н.Васильев, В.В.Михайлин

Предисловие

В учебном пособии отражены как теоретические, так и экспериментальные аспекты спектроскопии твердого тела. Пособие написано на основе курсов лекций «Физика конденсированного состояния вещества», «Спектроскопия твердого тела» и «Вторичные процессы в диэлектрических кристаллах», читаемых авторами студентам старших курсов кафедры оптики и спектроскопии физического факультета Московского государственного университета имени М.В.Ломоносова. Хотя по отдельным вопросам спектроскопии твердого тела есть много хороших книг, мы сделали попытку изложить вопросы оптической спектроскопии твердого тела в пособии небольшого объема. Мы ставили перед собой задачу дать в руки экспериментаторам, занимающимся спектроскопией твердого тела, пособие, которое бы включало основной минимум современных знаний, необходимый для этих занятий, и не требовало бы от читателя иной подготовки, кроме знаний общих курсов, читаемых всем студентам-физикам.

Спектроскопия твердого тела — достаточно старая и одновременно молодая область физики. С одной стороны, феноменологические законы, используемые в спектроскопии твердого тела, установлены очень давно (например, закон затухания света в веществе — закон Бугера—Ламберта, формулы Френеля для отражения света и др.) и являются той основой, без которой немислимо создание ни одного оптического прибора. С другой стороны, до сих пор не разрешены многие проблемы взаимодействия света с веществом, особенно в области вакуумного ультрафиолета, в которой источники излучения появились в основном лишь в последней четверти прошлого века. Предлагаемое издание и посвящено многим современным проблемам оптики твердых тел.

При изучении спектральных свойств твердого тела приходится рассматривать практически все аспекты физики твердого тела. Наиболее полное понимание процессов,

происходящих при взаимодействии электромагнитного излучения с веществом, стало возможным благодаря развитию квантовой теории твердого тела. Необходимо вспомнить о том, что одно из важнейших опытных обоснований квантовой теории — законы фотоэффекта — было изучено Эйнштейном именно как результат взаимодействия света с поверхностью твердого тела. Развитие теории твердого тела и выяснению микроскопических свойств твердых тел были посвящены многие работы Л.Д.Ландау, И.Е.Тамма, Я.И.Френкеля, А.Ф.Иоффе, И.М.Лифшица, В.Л.Гинзбурга, А.А.Абрикосова, Ж.И.Алферова и других ведущих советских и российских физиков. Их основополагающие исследования являются базисом, на котором строятся многие представления о твердом теле. И в настоящее время школа российских физиков вносит существенный вклад в исследования конденсированных состояний.

С развитием представлений о твердом теле меняется и предмет исследования спектроскопии твердого тела. Первоначально исследовались свойства твердых тел, описываемые феноменологическими постоянными, т. е. преломление, отражение света, дисперсия света в твердых телах и др. Теория электронной структуры твердых тел развилась, в частности, из оптических исследований, а с другой стороны, привела к более глубокому исследованию оптики твердых тел. В рамках этой теории описано издавна известное явление люминесценции. Однако только развитие практических применений люминесценции в 30-х годах заставило углубленно изучать этот интересный пример многочисленных вторичных процессов, возникающих в твердом теле после поглощения квантов электромагнитного излучения. Важную роль в этом сыграли работы С.И.Вавилова, В.Л.Лёвшина, других исследователей. Развитие исследования физических аспектов люминесценции привело к образованию крупных научных коллективов, занимающих ведущие позиции в этих вопросах. Из люминесценции выросло научное направление, производящее в настоящее время научно-технический переворот во многих областях человеческого знания и практики, — лазерная физика. Сейчас исследователей в области спектроскопии твердого тела занимают многие очень тонкие вопросы, позволяющие эффективно создавать качественно новые приборы, существенно расширяющие возможности человеческой практики и научных исследований.

В исследованиях по спектроскопии твердых тел одно из основных мест занимают источники излучений. Классические источники — лампы накаливания, низкотемпературные разрядные лампы — не в состоянии удовлетворить всем требованиям, предъявляемым к источникам света, в частности из-за ограниченности диапазона излучаемых длин волн. К 60-м годам созддалось такое положение, что существовало большое количество типов источников света в инфракрасной и видимой областях, ультрафиолете, а также в рентгене. Однако большинство источников в далеком ультрафиолете и рентгене давало лишь набор отдельных линий, что, естественно, не могло удовлетворить исследователей. В этом плане было особо ценным открытие и исследование, а затем и применение синхротронного излучения, позволившее за счет своего непрерывного спектра перекрыть область между ультрафиолетом и далеким рентгеном. В предсказании, развитии теории и исследовании синхротронного излучения основную роль сыграли советские физики Д.Д.Иваненко, И.Я.Померанчук, А.А.Соколов, И.М.Тернов, Ф.А.Королев. Синхротронное излучение, как оказалось, имеет и другие интересные свойства, делающие его уникальным для целого ряда исследований.

В настоящее время спектроскопия твердого тела охватывает настолько широкий круг вопросов, что осветить их все в одной книге практически невозможно. В связи с этим отбор материала, естественно, диктовался во многом интересами авторов. Это оправдано тем, что в настоящее время издано большое число превосходных книг по физике твердого тела, необходимо охватывающих и отдельные вопросы спектроскопии твердого тела [1—8]. Имеются и книги, посвященные только спектроскопии [9—18]. Однако в большинстве этих монографий и учебников рассматриваются вопросы оптики в традиционных областях спектра—от инфракрасного света до ультрафиолета, поэтому и объекты рассмотрения — это в основном металлы и полупроводники. С продвижением в область высокоэнергичных фотонов, в область вакуумного ультрафиолета появилась необходимость рассмотреть с единых позиций оптику широкозонных диэлектриков, имеющую свои особенности.

В одной книге трудно охватить все вопросы спектроскопии твердого тела, поэтому следует оговорить, что больше внимания мы уделяем диэлектрикам, чем полупроводникам и металлам. При описании техники измерений мы подробнее

описываем спектроскопию с применением синхротронного излучения, чем с лабораторными источниками. Несколько большее внимание уделено люминесценции и вторичным процессам вообще, возникающим в твердом теле при поглощении света. Кроме того, современные интересы спектроскопии переходят в несколько нетрадиционную область вакуумного ультрафиолета, в которой происходят переходы в области фундаментального поглощения. При этом проявляются все аспекты электронной структуры диэлектриков. Этой области спектра также уделено большое внимание. Список литературы включает в себя лишь основные книги, обзоры и статьи, которые могут помочь читателю усвоить материал книги.

Пособие состоит из нескольких частей. В первой части рассматриваются как общие подходы к описанию взаимодействия оптического излучения с конденсированными средами, так и элементарные электронные и фононные возбуждения. Во второй части пособия рассматриваются вопросы взаимодействия этих элементарных возбуждений, приводящие к релаксации энергии в твердых телах и появлению многочисленных вторичных процессов, возникающих при взаимодействии света с твердым телом - люминесценции, фотоэмиссии и проч.

Авторы выражают благодарность сотрудникам лаборатории синхротронного излучения кафедры оптики и спектроскопии МГУ за помощь в подготовке материала.

А.Н.Васильев, В.В.Михайлин,
октябрь 2008

Глава 1

Взаимодействие электромагнитного поля с твердыми телами

§ 1 Феноменологическое описание взаимодействия света с веществом

При изучении взаимодействия электромагнитного излучения с конденсированными состояниями вещества роль феноменологического описания трудно переоценить. Это справедливо не только для описания воздействия внешних электромагнитных полей на среду, но и для описания свойств фундаментальных возбуждений конденсированной среды. Поэтому вначале сформулируем основные понятия, характеризующие электромагнитные поля в вакууме и в среде.

1.1 Оптические функции твердых тел

Произвольное электромагнитное поле может быть полностью описано вектор-потенциалом $\mathbf{A}(\mathbf{r}, t)$ и скалярным потенциалом $\varphi(\mathbf{r}, t)$, действующими на точку \mathbf{r} в момент времени t . Эти потенциалы определяют электрическое поле $\mathbf{E}(\mathbf{r}, t)$ и магнитную индукцию $\mathbf{B}(\mathbf{r}, t)$ следующим образом (см., например, [19, 20]):

$$\mathbf{E}(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} - \nabla \varphi(\mathbf{r}, t),$$

$$\mathbf{B}(\mathbf{r}, t) = \text{rot } \mathbf{A}(\mathbf{r}, t).$$

Во многих случаях в пространственно-однородных системах удобно работать с монохроматическими или квази-монохроматическими плоскими полями. Из-за калибровочной инвариантности потенциалов скалярный потенциал $\varphi(\mathbf{r}, t)$ для таких полей часто можно выбирать равным нулю. Временная и пространственная зависимость вектор-потенциала характеризуется круговой частотой электромагнитного поля $\omega = 2\pi\nu$ и волновым вектором \mathbf{k} , например

$$\mathbf{A}(\mathbf{r}, t) = \text{Re}\left(\tilde{\mathbf{A}} e^{-i\omega t + i\mathbf{k}\mathbf{r}}\right).$$

Отсюда легко получить выражения для полей \mathbf{E} и \mathbf{B} . Здесь и ниже поведение комплексной плоской волны во времени описывается в форме $e^{-i\omega t}$, что определяет знаки мнимых частей всех вводимых ниже восприимчивостей. Комплексные величины обозначаются знаком '~' над буквой.

Плоские волны образуют полную систему функций, поэтому произвольное электромагнитное поле может быть представлено интегралом Фурье по плоским волнам:

$$\mathbf{A}(\mathbf{r}, t) = (2\pi)^{-4} \int d\omega d^3\mathbf{k} \tilde{\mathbf{A}}(\omega, \mathbf{k}) e^{-i\omega t + i\mathbf{k}\mathbf{r}}. \quad (1.1)$$

Если взаимодействие электромагнитного излучения с веществом может быть описано в линейном приближении, то для решения задачи с полем, имеющим произвольную временную и пространственную зависимости, необходимо знать только отклик системы на плоские монохроматические электромагнитные волны. Амплитуды плоских волн в формуле (1.1) выражаются в виде обратного преобразования Фурье:

$$\tilde{\mathbf{A}}(\omega, \mathbf{k}) = \int dt d^3\mathbf{r} \mathbf{A}(\mathbf{r}, t) e^{i\omega t - i\mathbf{k}\mathbf{r}}, \quad (1.2)$$

где интегрирование проводится по всему пространству и времени. Поскольку электромагнитные поля являются действительными,

$$\tilde{\mathbf{A}}(\omega, \mathbf{k}) = \tilde{\mathbf{A}}^*(-\omega, -\mathbf{k}).$$

При исследовании взаимодействия света с веществом важнейшую роль играет понятие частотной дисперсии, то есть зависимости отклика от частоты падающего излучения. Подавляющая часть спектроскопических исследований ограничивается изучением зависимости различных процессов от ω .

Поскольку значение ω для видимого света порядка $4 \cdot 10^{15}$ сек $^{-1}$ (для зеленого света) с соответствующим значением ν порядка $6 \cdot 10^{14}$ Гц, более удобно использовать длину волны $\lambda = c/\nu = 2\pi c/\omega$, выражаемую в нанометрах, или энергию фотона $h\nu = \hbar\omega$ в электрон-вольтах. Здесь $\hbar = h/2\pi = 1,055 \cdot 10^{-34}$ Дж·сек – постоянная Планка, а c – скорость света в вакууме. Энергия фотона связана с длиной волны соотношением

$$h\nu = 1239,9[\text{эВ} \cdot \text{нм}]/\lambda .$$

Другой единицей, применяемой для характеристики электромагнитных волн, являются обратные сантиметры.

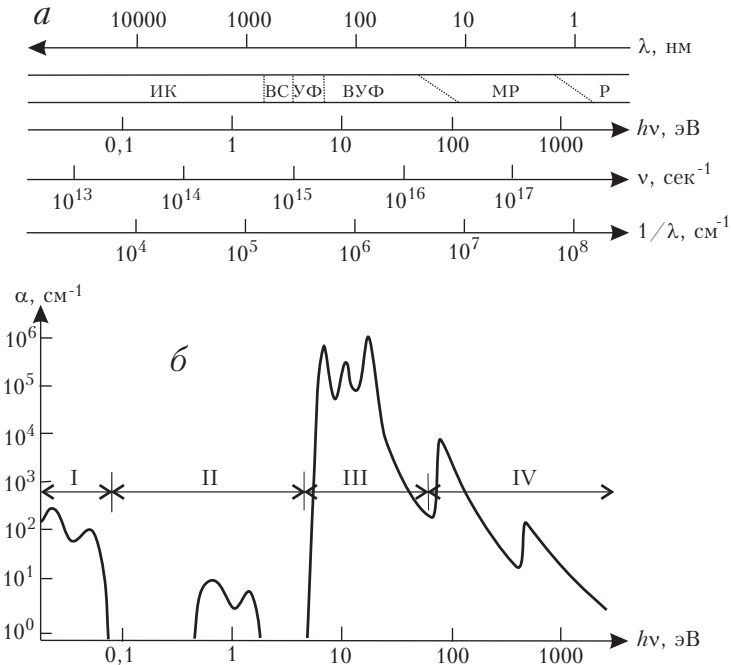


Рис. 1. Шкала электромагнитных волн (а) (ИК – инфракрасный свет, ВС – видимый свет, УФ – ультрафиолетовый свет, ВУФ – вакуумный ультрафиолетовый свет, МР – мягкий рентген, Р – рентген) и характерный спектр коэффициента поглощения (б) (I – область поглощения фононами, II – область примесного поглощения, III – область фундаментального поглощения, IV – область остовных переходов и рентгеновских краев поглощения)

Может быть полезным соотношение между электрон-вольтами и обратными сантиметрами:

$$1/\lambda = 8065,5 \left[\text{см}^{-1}/\text{эВ} \right] h\nu.$$

Эта единица измерения применяется в основном в инфракрасной спектроскопии или при изучении узких спектральных линий, например, в комбинационном рассеянии. Зеленый свет во введенных выше единицах измерения будет характеризоваться следующими значениями: $\lambda \approx 500 \text{ нм}$, $h\nu \approx 2,5 \text{ эВ}$, $1/\lambda \approx 20\,000 \text{ см}^{-1}$. На рис. 1а показана шкала электромагнитных волн, построенная в этих единицах. Границы между различными областями спектра в большой степени условны (кроме области видимого света, которая лежит от 770 нм (1,6 эВ) для красного цвета до 390 нм (3.2 эВ) для фиолетового цвета). В данной книге основное внимание будет уделяться вакуумной ультрафиолетовой (ВУФ) области спектра, то есть излучению, которое не проходит через воздух. Исследования в этой области требуют достаточно сложного вакуумного оборудования. Однако спектроскопические исследования диэлектриков с широкой запрещенной зоной дают большой объем информации именно в этой области.

Простейшим случаем является идеализированный случай взаимодействия плоской волны с однородной бесконечной средой. Электромагнитное поле создается поляризацией среды $\mathbf{P}(\mathbf{r}, t)$, пропорциональной разности электрической индукции $\mathbf{D}(\mathbf{r}, t)$ и $\mathbf{E}(\mathbf{r}, t)$. Используя плоские волны, это соотношение можно записать как

$$\tilde{\mathbf{D}}(\omega, \mathbf{k}) = \tilde{\mathbf{E}}(\omega, \mathbf{k}) + 4\pi\tilde{\mathbf{P}}(\omega, \mathbf{k}) = \tilde{\epsilon}(\omega, \mathbf{k})\tilde{\mathbf{E}}(\omega, \mathbf{k}) \quad (1.3)$$

Эта формула определяет диэлектрическую проницаемость среды $\tilde{\epsilon}(\omega, \mathbf{k})$, которая в общем случае является комплексным тензором, но для изотропной среды он сводится к скалярной комплексной функции. Для простоты мы здесь будем предполагать однородность и изотропность среды; случай кристаллов, для которых не применимо ни одно из этих предположений, будет описан ниже. Также мы не будем рассматривать магнитооптические эффекты. В связи с этим мы ограничимся немагнитными средами, для которых магнитная проницаемость $\tilde{\mu}$ равна единице.

Рассмотрим действие некоторого внешнего тока $\mathbf{j}_0(\mathbf{r}, t)$ на

систему, характеризующуюся диэлектрической проницаемостью $\tilde{\epsilon}(\omega, \mathbf{k})$. Фурье-компонента внешнего тока $\tilde{\mathbf{j}}_0(\omega, \mathbf{k})$ связана с $\mathbf{j}_0(\mathbf{r}, t)$ формулами, аналогичными формулам (1.1) и (1.2). Согласно уравнению непрерывности, внешний ток связан с изменением внешнего заряда $\rho_0(\mathbf{r}, t)$:

$$\partial \rho_0(\mathbf{r}, t) / \partial t + \operatorname{div} \mathbf{j}_0(\mathbf{r}, t) = 0 ,$$

что может быть записано через Фурье-компоненты

$$-i\omega \tilde{\rho}_0(\omega, \mathbf{k}) + i\mathbf{k} \tilde{\mathbf{j}}_0(\omega, \mathbf{k}) = 0 . \quad (1.4)$$

Таким образом, переменная часть внешнего заряда полностью определяется внешним током.

Уравнения Максвелла в среде

$$\begin{aligned} \operatorname{rot} \mathbf{H} &= \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c} \mathbf{j}_0, & \operatorname{div} \mathbf{D} &= 4\pi \rho_0, \\ \operatorname{rot} \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, & \operatorname{div} \mathbf{B} &= 0. \end{aligned}$$

могут быть переписаны с использованием Фурье-компонент в виде

$$\begin{aligned} \tilde{\mathbf{D}} &= -\frac{c}{\omega} [\mathbf{k} \tilde{\mathbf{H}}] - \frac{4\pi}{i\omega} \tilde{\mathbf{j}}_0, & \mathbf{k} \tilde{\mathbf{D}} &= 4\pi \tilde{\rho}_0, \\ \tilde{\mathbf{B}} &= \frac{c}{\omega} [\mathbf{k} \tilde{\mathbf{E}}], & \mathbf{k} \tilde{\mathbf{B}} &= 0. \end{aligned}$$

Отсюда в предположении изотропности среды, характеризующейся диэлектрической проницаемостью $\tilde{\epsilon}(\omega, \mathbf{k})$, можно получить уравнение для Фурье-компоненты электрического поля $\tilde{\mathbf{E}}(\omega, \mathbf{k})$:

$$\begin{aligned} \frac{c^2}{\omega^2} \left\{ \mathbf{k} (\mathbf{k} \tilde{\mathbf{E}}(\omega, \mathbf{k})) - k^2 \tilde{\mathbf{E}}(\omega, \mathbf{k}) \right\} + \tilde{\epsilon}(\omega, \mathbf{k}) \tilde{\mathbf{E}}(\omega, \mathbf{k}) &= \\ &= \frac{4\pi}{i\omega} \tilde{\mathbf{j}}_0(\omega, \mathbf{k}). \end{aligned} \quad (1.5)$$

Легко видеть, что фигурная скобка обращается в ноль, если $\tilde{\mathbf{E}}$ направлено вдоль вектора \mathbf{k} . Это дает возможность решить уравнение (1.5) относительно электрического поля $\tilde{\mathbf{E}}(\omega, \mathbf{k})$:

$$\tilde{\mathbf{E}}(\omega, \mathbf{k}) = \frac{4\pi}{i\omega} \left\{ \frac{\mathbf{k}(\mathbf{k}\tilde{\mathbf{j}}_0(\omega, \mathbf{k}))}{k^2\tilde{\epsilon}(\omega, \mathbf{k})} + \frac{k^2\tilde{\mathbf{j}}_0(\omega, \mathbf{k}) - \mathbf{k}(\mathbf{k}\tilde{\mathbf{j}}_0(\omega, \mathbf{k}))}{k^2[\tilde{\epsilon}(\omega, \mathbf{k}) - c^2k^2/\omega^2]} \right\}.$$

Это соотношение может быть переписано в виде, который лучше подчеркивает продольный и поперечный характер двух слагаемых решения:

$$\begin{aligned} \tilde{\mathbf{E}}(\omega, \mathbf{k}) &= \frac{4\pi\mathbf{k}\tilde{\rho}_0(\omega, \mathbf{k})}{ik^2\tilde{\epsilon}(\omega, \mathbf{k})} + \frac{4\pi[\mathbf{k}[\tilde{\mathbf{j}}_0(\omega, \mathbf{k})\mathbf{k}]]}{i\omega k^2[\tilde{\epsilon}(\omega, \mathbf{k}) - c^2k^2/\omega^2]} \equiv \quad (1.6) \\ &\equiv \tilde{\mathbf{E}}_l + \tilde{\mathbf{E}}_t \end{aligned}$$

Видно, что первый член этого уравнения соответствует продольному полю $\tilde{\mathbf{E}}_l(\omega, \mathbf{k}) \parallel \mathbf{k}$ и определяется только изменениями заряда, а второй соответствует поперечному полю $\tilde{\mathbf{E}}_t(\omega, \mathbf{k}) \perp \mathbf{k}$ и связан с изменениями компоненты тока, перпендикулярной волновому вектору \mathbf{k} .

Часто удобно выражать электрическое поле не через ток $\tilde{\mathbf{j}}_0$, а через поляризацию $\tilde{\mathbf{P}}_0$, связанную с током соотношением $\tilde{\mathbf{j}}_0(\mathbf{r}, t) = \partial\mathbf{P}_0(\mathbf{r}, t)/\partial t$ (или, в Фурье-компонентах, $\tilde{\mathbf{j}}_0(\omega, \mathbf{k}) = -i\omega\tilde{\mathbf{P}}_0(\omega, \mathbf{k})$):

$$\begin{aligned} \tilde{E}_i(\omega, \mathbf{k}) &= \\ &= 4\pi \left\{ \frac{k_i k_j}{k^2\tilde{\epsilon}(\omega, \mathbf{k})} + \frac{k^2\delta_{ij} - k_i k_j}{k^2[\tilde{\epsilon}(\omega, \mathbf{k}) - c^2k^2/\omega^2]} \right\} P_{0j}(\omega, \mathbf{k}) \quad (1.7) \end{aligned}$$

(по повторяющемуся индексу здесь и далее проводится суммирование).

Производя обратное Фурье-преобразование по волновому вектору \mathbf{k} , можно получить взаимосвязь электрического поля и поляризации в координатном представлении:

$$\tilde{E}_i(\omega, \mathbf{r}) = \int d^3r' \tilde{K}_{ij}(\mathbf{r} - \mathbf{r}') \tilde{P}_{0j}(\omega, \mathbf{r}').$$

Ядро \tilde{K}_{ij} может быть вычислено явным образом, если зависимость диэлектрической проницаемости $\tilde{\epsilon}$ от \mathbf{k} является слабой и ею можно пренебречь. В этом случае

$$\begin{aligned}
\tilde{K}_{ij}(\mathbf{r}) &= \\
&= \frac{4\pi}{(2\pi)^3} \int d^3k e^{i\mathbf{k}\mathbf{r}} \left\{ \frac{k_i k_j}{k^2 \tilde{\epsilon}(\omega)} + \frac{k^2 \delta_{ij} - k_i k_j}{k^2 [\tilde{\epsilon}(\omega) - c^2 k^2 / \omega^2]} \right\} = \\
&= \left\{ \frac{\omega^2}{c^2} \frac{\delta_{ij} r^2 - r_i r_j}{r^3} + \frac{3r_i r_j - \delta_{ij} r^2}{\tilde{\epsilon}(\omega) r^5} \left[1 + \sqrt{-\tilde{\epsilon}(\omega)} \frac{\omega r}{c} \right] \right\} \times \\
&\quad \times e^{-\sqrt{-\tilde{\epsilon}(\omega)} \omega r / c}.
\end{aligned} \tag{1.8}$$

В этой формуле знак квадратного корня из диэлектрической проницаемости определяется так, чтобы выполнялось условие $\text{Re} \sqrt{-\tilde{\epsilon}(\omega)} > 0$. В формуле (1.8) первый член в фигурных скобках соответствует излучению электромагнитных волн единичным диполем, помещенным в начало координат в среде с комплексной диэлектрической проницаемостью $\tilde{\epsilon}(\omega)$. Этот член в средах без поглощения спадает как $1/r$ и создает ненулевой поток электромагнитной энергии через сферу бесконечного радиуса.

Поскольку в пространственную зависимость поля (1.8) входит корень квадратный из диэлектрической проницаемости, то удобно ввести комплексный показатель преломления $\tilde{n} = n + ik$, связанный с диэлектрической проницаемостью $\tilde{\epsilon}$ формулой

$$\tilde{n}(\omega, \mathbf{k}) = \sqrt{\tilde{\epsilon}(\omega, \mathbf{k})}.$$

Действительная и мнимая части $\tilde{\epsilon} = \epsilon_1 + i\epsilon_2$ связаны с показателем преломления $n(\omega, \mathbf{k})$ и экстинкции $\kappa(\omega, \mathbf{k})$ следующим образом:

$$\epsilon_1 = n^2 - \kappa^2, \quad \epsilon_2 = 2n\kappa, \tag{1.9}$$

$$n = \sqrt{\frac{\epsilon_1 + \sqrt{\epsilon_1^2 + \epsilon_2^2}}{2}}, \tag{1.10}$$

$$\kappa = \sqrt{\frac{-\epsilon_1 + \sqrt{\epsilon_1^2 + \epsilon_2^2}}{2}}. \tag{1.11}$$

Причина того, что вводится две пары однозначно

связанных друг с другом функций, состоит в том, что в эксперименте измеряются значения n и κ , а теоретические расчеты наиболее прямо определяют ε_1 и ε_2 через микроскопические параметры.

При наличии поглощения ($\kappa(\omega) \neq 0$) уменьшение поля излучения определяется комбинацией множителя $1/r$ и

$$\exp[-in(\omega)\omega r/c - \kappa(\omega)\omega r/c] .$$

Второй член в фигурной скобке уравнения (1.8) описывает поле в ближней зоне и с ростом расстояния от диполя падает достаточно быстро. По этой причине в длинноволновом пределе можно пренебречь запаздыванием и положить в этом члене $\omega r/c = 0$, после чего ядро \tilde{K}_{ij} разбивается на сумму квазистационарного поля диполя и поля его электромагнитного излучения:

$$\begin{aligned} \tilde{K}_{ij}(\mathbf{r}) &= \\ &= \frac{3r_i r_j - \delta_{ij} r^2}{\tilde{\varepsilon}(\omega) r^5} + \frac{\omega^2}{c^2} \frac{\delta_{ij} r^2 - r_i r_j}{r^3} e^{-\kappa(\omega)\omega r/c - in(\omega)\omega r/c} . \end{aligned} \quad (1.12)$$

Поле излучения полностью поперечно (электрический вектор перпендикулярен радиус-вектору \mathbf{r} , направление которого совпадает с направлением распространения излучения), в то время как квазистационарное поле дипольного момента включает в себя как поперечные, так и продольные компоненты.

Из уравнения (1.6) видно, что электрическое поле резко возрастает, если знаменатель $\tilde{\varepsilon}(\omega, \mathbf{k})$ приближается к нулю. В случае

$$\tilde{\varepsilon}(\omega, \mathbf{k}) = 0 \quad (1.13)$$

продольное поле \mathbf{E}_l может существовать и без внешнего источника $\tilde{\rho}_0(\omega, \mathbf{k})$, то есть уравнение (1.13) является уравнением для свободных продольных колебаний. Соответственно, свободные поперечные колебания, существующие без внешнего источника тока $\tilde{\mathbf{j}}_0(\omega, \mathbf{k})$, определяются уравнением

$$\tilde{\varepsilon}(\omega, \mathbf{k}) = c^2 k^2 / \omega^2 . \quad (1.14)$$

Свет в вакууме представляет из себя поперечную электромагнитную волну, в которой векторы \mathbf{E} и \mathbf{B} перпендикулярны как друг другу, так и направлению распространения волны,

которое характеризуется единичным вектором \mathbf{s} . Волны с продольной поляризацией (в которых вектор \mathbf{E} параллелен направлению распространения) не существуют в вакууме (для которого $\tilde{\epsilon}(\omega, \mathbf{k}) = 1$). В то же время в среде с диэлектрической проницаемостью $\tilde{\epsilon}(\omega, \mathbf{k}) \neq 1$ могут существовать волны обоих типов.

Тем не менее понятие продольных электромагнитных полей вводится и в вакууме (например, при описании Фурье-разложения стационарного кулоновского поля). Такие поля не распространяются волновым образом, напротив, они локализованы в пространстве. Это связано с тем, что никакие действительные значения ω и \mathbf{k} , соответствующие фурье-компонентам из полного набора функций, по которым производится разложение произвольного поля, могут не удовлетворять секулярному уравнению (1.13).

Разложения (1.1) и (1.2) справедливы не только для изменяющихся во времени электромагнитных полей, но также и для стационарных и квазистационарных полей (например, кулоновских полей покоящихся или относительно медленно движущихся зарядов). Частота ω для таких полей равна нулю (точнее, много меньше частот, характерных для видимого света).

Выше уже было упомянуто, что диэлектрическая проницаемость $\tilde{\epsilon}(\omega, \mathbf{k}) = \epsilon_1(\omega, \mathbf{k}) + i\epsilon_2(\omega, \mathbf{k})$ является комплексной функцией. Это является фундаментальным свойством, связанным с принципом причинности (отличие от единицы $\epsilon_1(\omega, \mathbf{k})$ для каких-либо частот ω означает, что где-то при других ω связанная с поглощением мнимая часть диэлектрической проницаемости $\epsilon_2(\omega, \mathbf{k})$ также отлична от нуля; подробности и причины этого рассмотрены в разделе 1.3). Поэтому решения уравнений (1.13) и (1.14) в общем случае также являются комплексными. Характеристическое уравнение (1.14) для поперечных волн может быть переписано так, чтобы определить связь волнового вектора \mathbf{k} с частотой ω :

$$\mathbf{k} = \frac{\omega}{c} \tilde{n}(\omega, \mathbf{k}) \mathbf{s}, \quad (1.15)$$

где \mathbf{s} является единичным вектором, параллельным вектору \mathbf{k} . В качестве основных можно рассмотреть два случая. В первом случае, задаваясь действительным значением ω , получаем комплексное значение $\tilde{\mathbf{k}}$, которое описывает стационарное возбуждение колебаний с частотой ω . Этому типу

возбуждения соответствуют электромагнитные волны, огибающая которых затухает в пространстве по закону $e^{-k_2 r}$. Такая задача соответствует граничной задаче, когда источники стационарного электромагнитного поля расположены вне области, в которой распространяются поля. Случай действительного \mathbf{k} и мнимой частоты $\tilde{\omega}$ описывает релаксацию во времени начального возмущения. Пространственно-неоднородное возмущение может быть в этом случае разложено по функциям действительного \mathbf{k} , и при этом комплексные частоты $\tilde{\omega}$ описывают затухание во времени с огибающей $e^{-\omega_2 t}$. Таким образом, комплексная диэлектрическая проницаемость $\tilde{\epsilon}(\omega, \mathbf{k})$, зависящая от частоты ω и волнового вектора \mathbf{k} , позволяет проводить рассмотрение как пространственной, так и временной эволюции возмущений любого типа в среде, которая находится близко от равновесного состояния. Для сильного возбуждения среды (например, в области радиационного трека) необходимо, кроме того, учитывать изменение диэлектрической проницаемости во времени и пространстве.

Диэлектрическая проницаемость также определяет флуктуационные свойства электромагнитных полей в диэлектриках. Вообще говоря, в соответствии с флуктуационно-диссипационной теоремой интенсивности поглощения и флуктуаций тесно взаимосвязаны друг с другом (см., например, [21], а также раздел 3.3). Это дает возможность описать процессы релаксации возбуждений в терминах взаимодействия с флуктуациями электромагнитного поля. Такой подход, последовательно развитый Ю.Л.Климонтовичем [22] для плазмоподобных сред, позволяет провести полуфеноменологическое исследование релаксации электронных возбуждений.

Помимо двух пар функций (ϵ_1 и ϵ_2 ; n и κ), другой важной характеристикой вещества является коэффициент поглощения света. Для простоты мы ограничимся волнами с плоской поляризацией, распространяющейся вдоль оси x , с электрическим вектором, параллельным оси y (вектор-потенциал также параллелен этой оси). Модуль вектора Умова – Пойнтинга, характеризующий поток энергии в электромагнитной волне, равен

$$|\mathbf{S}| = S_x = \frac{c}{4\pi} [\mathbf{E}\mathbf{H}]_x = \frac{\omega^2 |\mathbf{A}|^2}{8\pi c} e^{-2\omega k x/c} . \quad (1.16)$$

Интенсивность плоской электромагнитной волны имеет такое же экспоненциальное затухание. Величина

$$\alpha = 2\omega\kappa/c = 4\pi\nu\kappa/c = 4\pi\kappa/\lambda \quad (1.17)$$

называется коэффициентом поглощения света в веществе. Интенсивность плоской волны на расстоянии x от поверхности описывается законом Ламберта $I \sim \exp(-\alpha x)$. Интенсивность электромагнитной волны, излучаемой единичным диполем (1.12), падает по формуле $\exp(-\alpha r)/r^2$. Коэффициент поглощения α чаще всего измеряется в обратных сантиметрах.

На рис. 16 схематически показан спектр поглощения диэлектрика, иллюстрирующий типичную частотную зависимость мнимой части диэлектрической проницаемости. Весь спектр можно разделить на четыре области, которые соответствуют различным типам возбужденных состояний в твердом теле. Детальный анализ этих областей составляет значительную часть содержания данной книги, поэтому здесь мы только кратко их опишем. Область I в большинстве диэлектриков попадает в область инфракрасного света (ИК), в которой свет взаимодействует с квантами колебаний решетки, называемых фононами. Коэффициент поглощения α в этой области достигает 10^2 см^{-1} . Вторая область соответствует области прозрачности кристалла. Область прозрачности в полупроводниках попадает в область ИК, в то время как для большинства оксидов и щелочно-галогидных кристаллов эта область тянется от ИК до ВУФ-диапазона. Из-за неидеальности кристаллов, в частности, примесей или дефектов, коэффициент поглощения α не равен нулю во всем диапазоне. В этой области наблюдаются полосы поглощения, для которых коэффициент поглощения α в максимумах достигает значений $10^{-1} - 10^2 \text{ см}^{-1}$, в зависимости от концентрации примесей или дефектов. Цвет кристаллов определяется именно этими полосами поглощения.

Область III называется областью фундаментального поглощения. Коэффициент поглощения резко возрастает до значений порядка $10^5 - 10^6 \text{ см}^{-1}$. Поглощение в этой области вызвано электронными переходами в атомах, составляющих регулярную решетку. Эти переходы ослабевают для энергии падающих фотонов порядка 20–40 эВ, и при этом падает коэффициент поглощения α . В области IV, которая соответствует еще большим энергиям (ультрамягкий рентген, УМР, и

рентгеновская область), наблюдаются новые скачки коэффициента поглощения, связанные с возбуждением из внутренних атомных оболочек.

Спектр на рис. 1 характеризует диэлектрики с широкой областью прозрачности. В полупроводниках область прозрачности намного уже, а в инфракрасной области наблюдается поглощение свободными носителями заряда. В металлах это поглощение доминирует в низкоэнергетичной области, в результате чего области прозрачности нет вообще. Разница между металлами и диэлектриками в областях III и IV выражена существенно слабее.

Описанная классификация различных областей всего спектра может быть проведена и при анализе других оптических функций. В каждой области спектра наиболее ярко проявляются возбужденные состояния кристалла определенного типа. Однако зачастую взаимодействие возбуждений друг с другом являются сильным, и только дополнительная информация из других областей спектра позволяет объяснить все свойства возбуждений.

Исследование диэлектрических констант α , ϵ_2 и проч. дает наиболее непосредственную информацию о внутренних процессах в твердых телах, связанных с взаимодействием со светом. Однако их прямое измерение зачастую затруднено. Например, для измерения α в области фундаментального поглощения требуются очень тонкие пленки (толщиной 10^{-5} - 10^{-4} см). По этой причине в спектроскопии широко применяются непрямые методы измерения диэлектрических констант. В частности, вместо измерения α часто производится измерение коэффициента отражения R , которое имеет и самостоятельное значение.

1.2 Отражение света от твердых тел

Для того чтобы получить связь коэффициента отражения с коэффициентами преломления n и экстинкции k , рассматривается идеализированная задача отражения света, падающего из вакуума на полубесконечную однородную среду под углом θ между нормалью к поверхности и направлением распространения света (рис. 2). При этом для получения коэффициента отражения падающей волны используют непрерывность тангенциальных к поверхности компонент электрического поля волны и скачок нормальных компонент с учетом комплексного показателя преломления \tilde{n} . Получаемые формулы носят

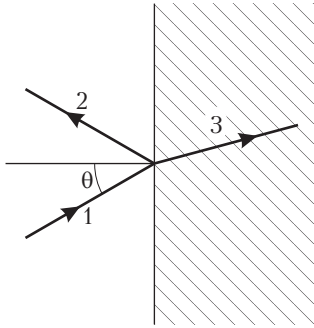


Рис. 2. Падение света на поверхность кристалла: 1 - падающий свет; 2 - отраженный; 3 - преломленный в кристалле свет

название формул Френеля (см., например, [23]). Поскольку в большинстве учебников их приводят для прозрачных сред, приведем их для поглощающих диэлектрических сред.

Комплексный коэффициент отражения \tilde{r}_{\perp} по амплитуде для падающей волны, поляризованной перпендикулярно плоскости падения, (электрический вектор лежит перпендикулярно плоскости падения, что обычно обозначается значком \perp или индексом s) равен

$$\tilde{r}_{\perp} = \frac{a^2 + b^2 - \cos^2 \theta - 2ib \cos \theta}{a^2 + b^2 + 2a \cos \theta + \cos^2 \theta}, \quad (1.18)$$

$$\text{где } a^2 = \frac{d + \sqrt{d^2 + 4n^2 \kappa^2}}{2}, \quad b^2 = \frac{-d + \sqrt{d^2 + 4n^2 \kappa^2}}{2},$$

$$d = n^2 - \kappa^2 - \sin^2 \theta.$$

При нормальном падении ($\theta = 0$) параметры a и b значительно упрощаются: $a = n$, $b = \kappa$.

Коэффициент отражения волн, поляризованной в плоскости падения (что обозначается знаком \parallel или индексом p), равен

$$\tilde{r}_{\parallel} = \tilde{r}_{\perp} \frac{a^2 + b^2 - \sin^2 \theta \operatorname{tg}^2 \theta + 2ib \sin \theta \operatorname{tg} \theta}{a^2 + b^2 + 2a \sin \theta \operatorname{tg} \theta + \sin^2 \theta \operatorname{tg}^2 \theta}. \quad (1.19)$$

Определяемый формулами (1.18) и (1.19) комплексный коэффициент отражения описывает изменения не только амплитуды, но и фазы электромагнитной волны. Однако в большинстве случаев измеряют только коэффициент отражения электромагнитной волны по интенсивности $R = |\tilde{r}|^2$, так что информация о фазе волны утрачивается. Для двух разных поляризаций коэффициенты отражения по интенсивности имеют вид

$$R_{\perp} = \frac{(a - \cos \theta)^2 + b^2}{(a + \cos \theta)^2 + b^2}, \quad R_{\parallel} = R_{\perp} \frac{(a - \sin \theta \operatorname{tg} \theta)^2 + b^2}{(a + \sin \theta \operatorname{tg} \theta)^2 + b^2}.$$

На рис. 3 приведена зависимость коэффициента отражения от угла падения θ . При $k = 0$ коэффициент отражения R_{\parallel} волны, поляризованной в плоскости падения, обращается в нуль для угла падения, равного углу Брюстера θ_{Br} . При малых $k \ll 1$ картина в целом такая же, но коэффициент отражения R_{\parallel} в нуль уже не обращается. При больших k зависимость коэффициента отражения от поляризации ослабевает. Коэффициент отражения не зависит от поляризации и для малых углов падения. Практически в области $0 \leq \theta < 12^\circ$ можно не различать поляризацию падающего излучения:

$$R = \left| \frac{\tilde{n} - 1}{\tilde{n} + 1} \right|^2.$$

При больших углах падения необходимо, однако, принимать во внимание как влияние поляризации, так и угла падения. Такая зависимость приводит к тому, что при применении зеркал в спектральной аппаратуре необходимо учитывать изменение поляризации при отражении. В области больших углов падения (скользящее падение, $\theta \rightarrow 90^\circ$) коэффициент отражения резко возрастает и приближается к единице. В области ультрамягкого рентгена и рентгена коэффициент отражения при нормальном падении для всех веществ существенно уменьшается, и для уменьшения потерь необходимо использовать отражательную оптику со скользящими углами падения. В случае $n < 1$ может наблюдаться хорошо известное явление полного внутреннего отражения (рис. 3б).

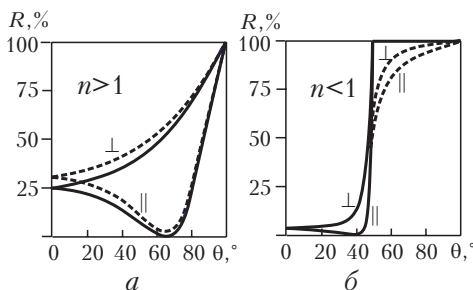


Рис. 3. Характерная зависимость коэффициента отражения света от угла падения для двух поляризаций в отсутствие поглощения (сплошные кривые) и при наличии слабого поглощения (пунктир)

Необходимо отметить, что слабое поглощение сильно влияет на коэффициент отражения в области углов больше угла полного внутреннего поглощения.

При изменении угла падения меняется и глубина проникновения света в вещество. Эффективный коэффициент поглощения оказывается равным

$$\alpha_{\text{eff}}(\theta) = \frac{2\omega}{c} b = \frac{\sqrt{2}\omega}{c} \sqrt{\left(\varepsilon_1 - \sin^2 \theta\right)^2 + \varepsilon_2^2} - \varepsilon_1 + \sin^2 \theta ,$$

и при наклонном падении света на поглощающее полупространство его интенсивность изменяется вдоль оси x , перпендикулярной плоскости поверхности, как $\exp(-\alpha_{\text{eff}}(\theta)x)$. В рентгеновской области спектра $\varepsilon_1 \approx 1$, $\varepsilon_2 \ll 1$, что позволяет упростить выражение для эффективного коэффициента поглощения:

$$\alpha_{\text{eff}}(\theta) = \alpha / \cos \theta .$$

На примере исследования отражения света можно проанализировать те сложности, которые встречаются на пути исследователей и вносят коррективы в интерпретацию экспериментов. Перечислим некоторые из них.

1. Как уже говорилось, в области фундаментального поглощения коэффициент α может достигать $10^5 - 10^6 \text{ см}^{-1}$. Это означает, что поглощение света происходит в слое толщиной порядка 10–100 нм — порядка сотни постоянных решеток. Это приводит к ряду ограничений на применение простых формул Френеля. Во-первых, поверхностный слой наиболее «грязный», он содержит существенно больше дефектов и примесей, чем основной объем кристалла. Особенно сильно поверхностный слой нарушается при механических воздействиях, например, полировке кристалла. Концентрация этих примесей может довольно сильно изменяться со временем за счет диффузии внутрь кристалла, взаимодействия с газовой средой, окружающей кристалл. Поэтому коэффициенты n и k могут существенно изменяться в приповерхностной области. К таким же последствиям приводит и изменение структуры даже чистого кристалла на поверхности. Во-вторых, поглощение в объеме отличается от поглощения вблизи поверхности еще и тем, что в последнем случае принимают участие и поверхностные электронные состояния (так называемые

таммовские состояния [24]; см. также [25]). Хотя плотность таммовских состояний невелика, однако их энергия может лежать в области прозрачности кристалла. Наконец, длина пробега электронных возбуждений l может быть достаточно большой, сравнимой с глубиной проникновения света в кристалл. Типичным примером этого явления может служить аномальный скин-эффект в металлах, когда l становится больше нормальной глубины скин-слоя, и электромагнитное поле «переносится» электронами вглубь кристалла. При этом меняется зависимость коэффициента отражения от частоты падающего света.

2. Аномальный скин-эффект является одним из примеров пространственной дисперсии показателей n и k — зависимости \tilde{n} от волнового вектора электромагнитной волны \mathbf{k} . При учете этой зависимости уравнение для определения волнового вектора электромагнитной волны в однородном кристалле (1.15) может иметь уже не одно решение. При этом в кристалле возможно распространение нескольких волн с одной и той же частотой ω , но разными (вообще говоря, комплексными) векторами $\tilde{\mathbf{k}}$. Помимо аномального скин-эффекта, к такого рода зависимостям могут приводить эффекты сильного взаимодействия между фотонами и элементарными возбуждениями в кристалле (поляритоны разных типов). Строгий путь исследования задачи отражения предусматривает решение микроскопических уравнений с соответствующими граничными условиями на функции распределения возбуждений. Однако возможен и феноменологический учет появления добавочных волн в кристалле. При сшивке решений вне и внутри кристалла, аналогичной сшивке при получении формул Френеля, используют так называемые дополнительные граничные условия. В настоящее время предложено довольно много различных формулировок дополнительных граничных условий, применяемых в конкретных задачах (см., например, [10]).
3. При измерениях в области прозрачности необходимо учитывать конечность толщины кристалла, т. е. принимать во внимание отражение от задней грани или от границы раздела кристалл — подложка. В случае, когда толщина кристалла (или пленки, выращенной на поверхности подложки) порядка длины волны, помимо отражения необходимо учитывать и интерференцию от задней грани.

4. Реально исследуемые поверхности не являются геометрически идеальными. Даже у кристаллов высокого качества на поверхности наблюдаются ступени роста. По мере увеличения энергии квантов падающего света высота этих ступеней становится сравнимой с длиной волны, и при этом необходимо учитывать возникновение интерференции при отражении. Еще одно следствие геометрической неидеальности поверхности — появление диффузной компоненты в отраженном свете, когда помимо отражения света под углом падения наблюдается рассеяние по всем направлениям. Этот эффект также усиливается при переходе к более коротковолновому излучению. Эффект диффузного отражения в полной мере проявляется в случае работы с поликристаллами и порошками, состоящими из большого числа монокристалликов разного размера. При этом наблюдаются многократное отражение света, прохождение уже отраженного света через микрокристаллы и т. д. Количественное описание этого явления возможно только в достаточно специфических случаях из-за множества параметров, описывающих порошок и влияющих на процессы взаимодействия света с порошком (соотношение длины волны, глубины проникновения α^{-1} и характерного размера частиц порошка, функции распределения частиц по размерам и ориентации граней микрокристаллов, отношения объема, занимаемого кристаллами порошка, к свободному объему между кристалликами и т. д.), поэтому использование диффузного отражения для количественного определения диэлектрических констант является чрезвычайно сложной задачей.
5. При выводе формул (1.18), (1.19) предполагалась изотропность вещества. В действительности для многих типов кристаллов необходимо учитывать эффекты анизотропии, когда свет распространяется с разной скоростью в разных направлениях в кристалле. При этом вместо скалярных величин ϵ_1 и ϵ_2 вводятся тензоры $(\epsilon_1)_{ij}$ и $(\epsilon_2)_{ij}$. Если главные оси этих тензоров совпадают и перпендикулярны одной из граней кристалла, то при исследовании отражения света, нормально падающего на эту грань, можно получить соответствующие диагональные элементы этих тензоров вдоль рассматриваемой оси по обычным формулам для изотропного кристалла. В противном случае требуется довольно громоздкий анализ [23].

6. При изучении спектров отражения необходимо иметь в виду, что при очень сильных интенсивностях падающего света (например, при исследованиях с помощью лазеров) возможно проявление нелинейных эффектов (умножение частоты, параметрические эффекты). При описании поглощения в этом случае уже нельзя пользоваться линейной теорией с восприимчивостями n и k .
7. В средах с искусственно созданной периодической структурной неоднородностью с характерными размерами порядка длины волны наблюдаются ярко выраженные дифракционные эффекты, проявляющиеся в сильной зависимости характеристик распространения света от направления и частоты (эффекты "фотонных кристаллов", неотражающие покрытия, так называемые среды "с отрицательным показателем преломления"). Эта тема в настоящее время привлекает большое внимание.
8. Наконец, процесс взаимодействия света с веществом сводится не просто к отражению света. Поглощение энергии падающих квантов света приводит к множеству вторичных процессов: рассеяние света в кристалле без изменения частоты комбинационное рассеяние с относительно слабым изменением частоты падающего света, люминесценция, в том числе и с довольно значительной временной задержкой, фотоэффект и пр. При этом в практических измерениях бывает необходимо отделять свет люминесценции от отраженного света и т. д. Из-за различия диаграмм направленности люминесценции и зеркально отраженного света, из-за частотных сдвигов разделение этих процессов в принципе несложно осуществить. В настоящей книге подробно рассмотрены многие вторичные процессы, поэтому здесь можно ограничиться сделанными замечаниями.

Таким образом, исследование даже такого простого процесса, как отражение света от кристаллов, связано с рядом принципиальных ограничений и усложнений. Все приведенные выше различные факторы влияют не только на процесс отражения света, но и на многие другие спектроскопические эффекты.

Рассмотрим коротко ряд способов измерения оптических констант. В широкой области спектра возможно непосредственное измерение только коэффициента поглощения α . Непосредственное измерение n ограничивается областью прозрачности кристалла ($\alpha \leq 1 \text{ см}^{-1}$). Однако измерения

поглощения в области фундаментального поглощения, как уже указывалось, требуют очень тонких слоев вещества. Поэтому большое число косвенных методов определения оптических констант предусматривает измерение коэффициента отражения. Поскольку из этих измерений необходимо получить две независимые константы n и k (или ε_1 и ε_2), то приходится измерять два раза коэффициент отражения света при разных условиях. Можно привести несколько таких способов.

1. Измерение коэффициентов отражения света R_1 и R_2 , падающего под двумя углами к поверхности, θ_1 и θ_2 .
2. Измерение коэффициентов R_1 и R_2 света с различной поляризацией.
3. Иммерсионные методы, при которых сравниваются коэффициенты отражения от кристалла в вакууме (на воздухе) и в прозрачной жидкой среде с известным показателем преломления. Эти методы применимы только в той области энергий падающих фотонов, в которых прозрачны жидкости.

Перечисленные методы приводят к системе двух уравнений, выражающих, например, R_1 и R_2 через константы n и k . В принципе решить такие системы возможно. Однако эти методы имеют следующее принципиальное ограничение. В ряде областей изменения n и k якобиан преобразования $\partial(R_1, R_2)/\partial(n, k)$ очень мал, т. е. малым изменениям коэффициентов отражения ΔR_1 , ΔR_2 соответствуют большие изменения оптических констант Δn и Δk . Поскольку точность измерения R ограничена, незначительные ошибки при определении коэффициента отражения приводят к недопустимым ошибкам при вычислении оптических констант. Это сильно сужает область применимости приведенных выше методов.

В ряде случаев эффективно применяются эллипсометрические измерения. При этом исследуется превращение линейной поляризации падающего света в эллиптическую, что дает возможность определить комплексный коэффициент отражения \tilde{r}_\perp или \tilde{r}_\parallel . Метод ограничен наличием сложностей получения поляризованного света в некоторых областях спектра.

1.3 Соотношения Крамерса—Кронига и правила сумм

Все перечисленные выше методы определения оптических констант требуют измерения двух величин на одной частоте падающего света. Существует достаточно надежный способ определения оптических констант по измерениям

коэффициента отражения света в широкой области спектра. Предварительно необходимо сделать несколько замечаний о свойствах оптических констант [19].

Диэлектрическая проницаемость $\tilde{\epsilon}(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega)$ связывает Фурье-компоненты индукции $\tilde{\mathbf{D}}(\omega)$ и напряженности $\tilde{\mathbf{E}}(\omega)$ электрического поля. Функцию $\tilde{\epsilon}(\omega)$, определенную при действительных ω , можно аналитически продолжить в область комплексных $\tilde{\omega}$. Аналитические свойства $\tilde{\epsilon}(\tilde{\omega})$ определяются принципом причинности, согласно которому индукция $\mathbf{D}(t)$ может быть связана с напряженностью электрического поля $\mathbf{E}(t)$ лишь в предшествующие моменты времени:

$$\mathbf{D}(t) = \int_{-\infty}^t dt' \hat{\epsilon}(t-t') \mathbf{E}(t) , \quad (1.20)$$

то есть обратное Фурье-преобразование $\hat{\epsilon}(t)$ от $\tilde{\epsilon}(\omega)$ обращается в ноль для отрицательных значений времен. Из теории функций комплексного переменного известно, что это возможно только в случае, если $\tilde{\epsilon}(\tilde{\omega})$ не будет иметь особенностей в верхней полуплоскости $\tilde{\omega}$. Применение этой теории к такого рода функции дает возможность записать интегральную связь ϵ_1 и ϵ_2 :

$$\epsilon_1(\omega) - 1 = \frac{1}{\pi} \text{V.p.} \int_{-\infty}^{\infty} \frac{\epsilon_2(x)}{x - \omega} dx , \quad (1.21)$$

$$\epsilon_2(\omega) - \frac{4\pi\sigma}{\omega} = -\frac{1}{\pi} \text{V.p.} \int_{-\infty}^{\infty} \frac{\epsilon_1(x) - 1}{x - \omega} dx . \quad (1.22)$$

Из-за наличия сингулярности в точке $x = \omega$ интегралы воспринимаются в смысле главного значения (что обозначается символом V.p.). Дополнительный член -1 в левой части (1.21) появляется из-за того, что при высоких частотах ω , много больших всех характерных частот движения электронов в веществе, $\tilde{\epsilon} \rightarrow 1$ (более подробно это поведение обсуждается ниже). Второй член в левой части уравнения (1.22) возникает в проводящих кристаллах с проводимостью σ , для которых диэлектрическая проницаемость при малых частотах имеет полюс $4\pi i\sigma/\omega$. При рассмотрении аналитических свойств $\tilde{\epsilon}(\tilde{\omega})$ этот полюс следует исключить явным образом. Для диэлектриков $\sigma = 0$. Случай сверхпроводящих

кристаллов, для которых $\varepsilon_1 \sim 1/\omega^2$ при $\omega \rightarrow 0$, мы рассматривать не будем.

Соотношения (1.21) и (1.22) по имени авторов называются дисперсионными соотношениями Крамерса – Кронига [26, 27]. Эти соотношения в принципе позволяют знать только одну из функций (ε_1 или ε_2), но во всей области частот, и получить по ним вторую функцию.

Частота ω в соотношениях (1.21) и (1.22) (равно как и переменная интегрирования x) может меняться в пределах от $-\infty$ до ∞ . Более естественно работать только с положительными значениями частоты. Для этого необходимо связать $\tilde{\varepsilon}(-\omega)$ с $\tilde{\varepsilon}(\omega)$. В формуле (1.20) $\hat{\varepsilon}(t-t')$ связывает две действительные величины $\mathbf{D}(t)$ и $\mathbf{E}(t)$, поэтому $\hat{\varepsilon}(t-t')$ тоже является действительной функцией. Отсюда следует, что ее Фурье-компоненты удовлетворяют соотношению $\tilde{\varepsilon}(-\omega) = \tilde{\varepsilon}^*(\omega)$ или

$$\varepsilon_1(\omega) = \varepsilon_1(-\omega), \quad \varepsilon_2(\omega) = -\varepsilon_2(-\omega). \quad (1.23)$$

Тогда формулы (1.21) и (1.22) могут быть приведены только к положительным частотам:

$$\varepsilon_1(\omega) - 1 = \frac{2}{\pi} \text{V.p.} \int_0^{\infty} \frac{x \varepsilon_2(x)}{x^2 - \omega^2} dx, \quad (1.24)$$

$$\varepsilon_2(\omega) - \frac{4\pi\sigma}{\omega} = -\frac{2\omega}{\pi} \text{V.p.} \int_{-\infty}^{\infty} \frac{\varepsilon_1(x) - 1}{x^2 - \omega^2} dx. \quad (1.25)$$

Соотношения (1.21) и (1.22) представляют и самостоятельный интерес, и мы еще обсудим их ниже. Вернемся теперь к первоначальной цели нашего обсуждения – получению оптических функций из спектров отражения. В принципе можно получить систему двух нелинейных интегральных уравнений, используя связь ε_1 и ε_2 с комплексным коэффициентом отражения \tilde{r} . Однако решить такую систему чрезвычайно трудно. Рядом авторов (см., например, [28, 29]) был предложен метод нахождения оптических констант из дисперсионных соотношений несколько другого вида. При этом в качестве аналитической функции, не имеющей полюсов в верхней полуплоскости комплексной частоты $\tilde{\omega}$, рассматривается логарифм комплексного коэффициента отражения по амплитуде $\tilde{r} = \sqrt{R} e^{i\varphi}$, где $\varphi(\omega)$ – фазовый сдвиг электромагнитной волны

после отражения от кристалла. Для функции $\ln \tilde{r}(\omega) = \frac{1}{2} \ln R(\omega) + i\varphi(\omega)$ можно записать

$$\varphi(\omega) = \frac{\omega}{\pi} \int_0^{\infty} \ln \frac{R(x)}{R(\omega)} \frac{dx}{x^2 - \omega^2}. \quad (1.26)$$

В этом соотношении нет необходимости рассматривать интеграл в смысле главного значения, поскольку $\ln[R(x)/R(\omega)]$ обращается в ноль при $x = \omega$, и при этом особенность в интеграле исчезает. В действительности $\ln \tilde{r}$ не является «хорошей» аналитической функцией в верхней полуплоскости, поскольку \tilde{r} может обращаться в ноль при некоторых значениях $\tilde{\omega}$ (например, для действительной ω и p — поляризации света — при угле Брюстера θ_{Br}), что проявляется в наличии точек ветвления $\ln \tilde{r}$. Поэтому равенство (1.26) имеет приближенный характер (более точно, соотношение (1.26) удовлетворяется с точностью до некоторой гладкой функции от частоты ω , поэтому форма особенностей воспроизводится правильно). Несмотря на это, соотношения Крамерса — Кронига для отражения часто применяются для получения оптических констант, особенно в области фундаментального поглощения диэлектриков.

Зная коэффициент отражения $R(\omega)$, можно вычислить фазу $\varphi(\omega)$, а затем и константы n и k . При этом приходится сталкиваться с тем обстоятельством, что коэффициент отражения известен только в ограниченной области частот — от ω_{\min} до ω_{\max} . Ядро интегрального преобразования (1.26) медленно спадает с ростом частоты, поэтому вклад внешних областей от 0 до ω_{\min} и от ω_{\max} до ∞ является существенным, и возникает задача аппроксимации R в двух внешних областях. Обычно коэффициент отражения в области малых частот можно аппроксимировать литературными данными. Кроме того, область до ω_{\min} обычно попадает на область прозрачности вещества, при этом $\varphi(\omega) = 0$, поскольку в этом диапазоне \tilde{r} — действительная величина (см. (1.18), (1.19)). Аппроксимация в высокочастотной области обычно основывается на поведении диэлектрической проницаемости на очень больших частотах. Если эти частоты много больше характерных частот электронов в атомах, то такие электроны могут рассматриваться как свободные и может применяться формула для диэлектрической проницаемости свободного электронного газа

[30]:

$$\varepsilon_1(\omega) = 1 - \frac{4\pi N e^2}{m\omega^2}, \quad \varepsilon_2(\omega) = o\left(1/\omega^3\right) \text{ при } \omega \rightarrow \infty. \quad (1.27)$$

Здесь N — число электронов в единице объема во всех атомах, m и e — масса и заряд свободного электрона. Используя эту зависимость, можно показать, что на больших частотах $R \sim \omega^{-4}$. К сожалению, границы применимости этой формулы лежат выше частот всех электронных переходов в твердом теле. Для тяжелых атомов это соответствует далекому рентгену. Однако, если ω_{\max} находится в области, в которой переходы из внутренних оболочек атомов еще невозможны, а переходы из внешних оболочек уже исчерпаны, диэлектрическая проницаемость также описывается формулой (1.27), в которой под величиной N понимается число электронов на соответствующих внешних оболочках атомов.

Кроме сделанных аппроксимаций, часто требуется дополнительная коррекция фазы после вычисления интеграла (1.26). Это связано как с неточностями примененных приближений и ошибками в экспериментальных данных, так и с отмеченным выше нестрогим характером вывода соотношения (1.26). Коррекция состоит в добавлении к рассчитанной фазе гладкой функции, которая выбирается исходя из следующих критериев:

- а) во всей области прозрачности вплоть до области фундаментального поглощения фаза $\varphi(\omega)$ обращается в ноль;
- б) в области малого коэффициента поглощения на больших частотах фаза $\varphi(\omega)$ приближается к π (по аналогии с колебаниями маятника под действием внешней силы, который колеблется в противофазе с внешней силой, если частота последней много больше резонансной частоты маятника);
- в) наилучшим критерием является сравнение результатов расчета с оптическими характеристиками, полученными в одной или нескольких точках другими методами; к сожалению, обычно трудно получить такую информацию;
- г) контроль результатов расчета по точности выполнения соотношений Крамерса — Кронига для функций ε_1 и ε_2 , вычисленных из $R(\omega)$ и $\varphi(\omega)$;
- д) контроль точности выполнения описанных ниже правил сумм.

Этот метод определения диэлектрических констант с

помощью соотношений Крамерса — Кронига является одним из основных в области фундаментального поглощения (рис. 4, 5).

Как правило, вычисления по формулам (1.24) — (1.26) проводятся как с использованием метода Симпсона с выделением из интеграла области вблизи точки сингулярности, так и с использованием специальных процедур интегрирования. Широко применяется линейная или квадратичная интерполяция экспериментальных данных, что приводит к появлению логарифмов под знаком сумм. Поэтому прямое вычисление по соотношениям Крамерса — Кронига для N точек требует порядка N^2 умножений или вычислений логарифма. Для ускорения расчета можно применять метод [31, 32], основанный на быстром преобразовании Фурье, что позволяет проводить вычисления с коррекцией в интерактивном режиме.

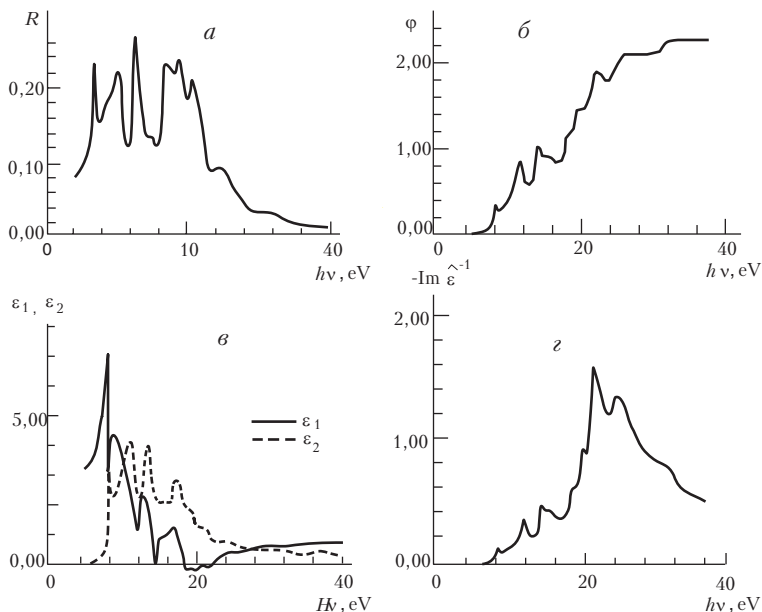


Рис. 4. Экспериментальный спектр отражения монокристалла MgO (а) [34] и вычисление оптических констант с помощью соотношений Крамерса — Кронига (фазовый угол φ (б), функции ϵ_1 , ϵ_2 (в) и $\text{Im}(-\hat{\epsilon}^{-1})$ (г))

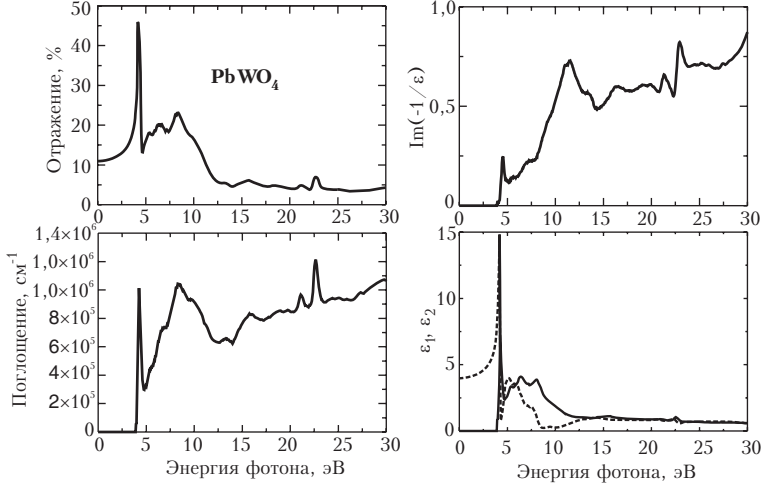


Рис. 5. Спектр отражения вольфрамата свинца [36] и пересчитанные с использованием БПФ оптические функции.

В математике преобразования (1.21) и (1.22) называются преобразованиями Гильберта (см., например, [33]). Преобразование Гильберта от $\sin \omega t$ равно $\cos \omega t$, а от $\cos \omega t$ равно $-\sin \omega t$. Если произвольная функция $f(\omega)$ представлена интегралом Фурье

$$f(\omega) = \int_{-\infty}^{\infty} [a(t) \cos \omega t + b(t) \sin \omega t] dt,$$

то ее преобразование Гильберта имеет вид

$$g(\omega) = \int_{-\infty}^{\infty} [b(t) \cos \omega t - a(t) \sin \omega t] dt.$$

Поэтому процедура вычислений преобразований Крамерса—Кронига сводится к последовательному применению прямого и обратного преобразований Фурье. Если применять алгоритм быстрого преобразования Фурье (БПФ), то потребуется только $N \ln N$ умножений.

Дисперсионные соотношения (1.21) и (1.22) позволяют получить ряд интересных интегральных соотношений (см.,

например, [35]). Как уже было отмечено выше, ε_2 на больших частотах стремится к нулю, и поэтому можно записать

$$\varepsilon_1(\omega) - 1 = -\frac{2}{\pi\omega^2} \int_0^{\infty} x\varepsilon_2(x) dx \text{ при } \omega \rightarrow \infty. \quad (1.28)$$

Сравнение (1.28) с соотношением (1.27) дает правило сумм в следующей форме:

$$\frac{m}{2\pi^2 e^2} \int_0^{\infty} \omega\varepsilon_2(\omega) d\omega \equiv \int_0^{\infty} f(\omega) d\omega = N. \quad (1.29)$$

Здесь введена сила осцилляторов $f(\omega) \equiv \frac{m\omega\varepsilon_2(\omega)}{2\pi^2 e^2} d\omega$, принадлежащих на интервал частот $d\omega$. Правило сумм (1.29) выражает тот факт, что сумма всех сил осцилляторов в единице объема равна числу электронов в единице объема. Интеграл в (1.29) должен быть сходящимся, откуда следует, что $\varepsilon_2(\omega)$ при больших частотах падает с ростом ω быстрее, чем ω^{-2} .

В другом пределе $\omega \rightarrow 0$ для диэлектриков ($\sigma = 0$) можно получить:

$$\varepsilon_1(0) - 1 = \frac{2}{\pi} \int_0^{\infty} \frac{\varepsilon_2(\omega)}{\omega} d\omega \equiv \frac{4\pi e^2 N}{m} \overline{\omega^{-2}}, \quad (1.30)$$

где черта означает усреднение с помощью силы осцилляторов:

$$\overline{\omega^{-2}} \equiv \frac{1}{N} \int_0^{\infty} \frac{1}{\omega^2} f(\omega) d\omega.$$

В уравнении (1.30) $\varepsilon_1(0)$ — статическая диэлектрическая проницаемость, обычно обозначаемая как ε_0 . Необходимо обратить внимание на то, что интегрирование в (1.30) производится по всему диапазону частот, в том числе и по области поглощения в ИК-диапазоне с участием фононов. Вообще говоря, в отличие от формулы (1.30), которая чувствительна к поведению $\varepsilon_2(\omega)$ при сравнительно низких частотах, правило сумм (1.29) в основном определяется поведением $\varepsilon_2(\omega)$ при высоких частотах. Оценка точности расчета по правилам Крамера — Кронига с помощью этих правил сумм помогает выбрать аппроксимацию для высоко- и низкочастотных частотных

диапазонов.

При анализе оптических функций с помощью правил сумм часто вводится эффективное число электронов $N_{\text{eff}}(\omega)$ и эффективная статическая диэлектрическая проницаемость $\epsilon_{\text{eff}}^0(\omega)$. Они получаются по формулам (1.29) и (1.30), в которых верхний предел интегрирования заменен на частоту ω . Эти две функции интерпретируются как число электронов, участвующих в переходах в диапазоне частот от 0 до ω и как вклад таких электронов в ϵ_0 . Быстрый ступенчатый рост таких функций указывает на начало переходов из все более глубоких оболочек. К сожалению, точное определение таких функций оказывается невозможным. Более того, они крайне чувствительны (особенно $N_{\text{eff}}(\omega)$) к ошибкам подгонки при вычислении преобразований Крамерса – Кронига.

Правила сумм, аналогичные соотношениям (1.29) и (1.30), можно записать и для $n(\omega)$ и $\kappa(\omega)$, а также для других оптических функций (например, используемой ниже функции потерь энергии $\text{Im}(-\epsilon^{-1}(\omega))$).

§ 2 Характеристические потери энергии электронов

При прохождении через вещество заряженная частица интенсивно взаимодействует со всеми возможными возбуждениями. Рассматривая только электростатическое взаимодействие через продольное электрическое поле, необходимо учитывать поляризацию вещества (см., напр. [37–39]). Можно показать, что потери энергии электрона на создание возбуждения с импульсом $\hbar\mathbf{q}$ и энергией $\hbar\omega$ пропорциональны

$$\text{Im}\left(-\frac{1}{\tilde{\epsilon}(\omega, \mathbf{q})}\right) = \frac{\epsilon_2(\omega, \mathbf{q})}{|\tilde{\epsilon}(\omega, \mathbf{q})|^2}. \quad (2.1)$$

Вообще говоря, $\tilde{\epsilon}(\omega, \mathbf{q})$ в этой формуле должна соответствовать продольной, а не поперечной части тензора диэлектрической проницаемости (более подробно различие между этими двумя тензорами рассматривается в следующей главе). Формула (2.1) связана с наличием $\tilde{\epsilon}^{-1}$ в выражении (1.6) для продольного электрического поля \mathbf{E}_l . Электрон (или другая заряженная частица) создает в кристалле поле с потенциалом $\phi(r) = -e/\hat{\epsilon}r$, где $\hat{\epsilon}$ - оператор диэлектрической проницаемости. Фурье-компонента этого поля имеет вид $4\pi e/q^2 \tilde{\epsilon}(\omega, \mathbf{q})$, где $\tilde{\epsilon}(\omega, \mathbf{q})$ - фурье образ оператора диэлектрической проницаемости по пространственным координатам и времени. Для оценки потерь энергии электрона в формуле (2.1) можно положить $\mathbf{q} = 0$ и проинтегрировать по всем частотам:

$$-\frac{dE}{dx} \sim \int d\omega \text{Im}\left(-\frac{1}{\tilde{\epsilon}(\omega)}\right).$$

Рассмотрим более подробно свойства функции $\text{Im}\left(-\tilde{\epsilon}(\omega)^{-1}\right)$. Во-первых, согласно (2.1), потери энергии

квантами $\hbar\omega$ отличны от нуля в двух случаях. Если $\varepsilon_2(\omega) \neq 0$, то возбуждаются такие электронные состояния, которые дают вклад в $\varepsilon_2(\omega)$ на соответствующей частоте. Тем самым потери энергии приводят к рождению новых электронных возбуждений. Правда, вклад разных состояний меняется в силу изменения $1/|\tilde{\varepsilon}(\omega)|^2$ с частотой. В частности, он возрастает с уменьшением $|\tilde{\varepsilon}(\omega)|^2$. Наиболее интересный вклад в потери энергии возможен и при $\varepsilon_2(\omega) = 0$. Если определить ω_{pl} как корень уравнения $\varepsilon_1(\omega_{\text{pl}}) = 0$, то

$$\frac{\varepsilon_2(\omega)}{|\tilde{\varepsilon}(\omega)|^2} \approx \frac{\pi\omega_{\text{pl}}}{2} [\delta(\omega - \omega_{\text{pl}}) - \delta(\omega + \omega_{\text{pl}})].$$

Частота, на которой ε_1 обращается в ноль, называется плазменной частотой. При этом возбуждаются особые квази-частицы — плазмоны, являющиеся продольными колебаниями электрического поля. Случай $\varepsilon_2(\omega) = 0$ соответствует модели свободных электронов. При этом

$$\varepsilon(\omega) = 1 - \frac{4\pi N e^2}{m\omega^2}, \quad (2.2)$$

где N — концентрация и m — масса свободных электронов. Из (2.2) получаем

$$\omega_{\text{pl}} = \sqrt{\frac{4\pi N e^2}{m}}. \quad (2.3)$$

Если поглощение отлично от нуля, $\varepsilon_2(\omega) \neq 0$, то все равно на больших частотах $\varepsilon_2(\omega) \rightarrow 0$. В этой области соотношение (2.2) связано с правилами сумм, следующими из дисперсионных соотношений Крамерса — Кронинга (1.24). В качестве N здесь стоит плотность валентных электронов, переходы с которых к ω_{pl} уже исчерпаны. Однако в большинстве случаев выражение для ω_{pl} (2.3) представляет собой лишь грубую оценку. Даже в металлах это соотношение в большинстве случаев не применимо. В действительности плазменную частоту необходимо определять из условия $\varepsilon_1(\omega_{\text{pl}}) = 0$. В силу комплексности $\tilde{\varepsilon}$ выражение для плазменной частоты также получается комплексным, причем мнимая часть $\tilde{\omega}_{\text{pl}}$ растет с

ростом $\varepsilon_2(\text{Re } \omega_{pl})$. Тем самым плазмоны становятся короткоживущими. В большинстве диэлектриков максимум на частоте ω_{pl} имеет полуширину порядка нескольких электрон-вольт (см., например, рис. 4–6), т.е. время жизни плазмона порядка 10^{-15} сек. Только в металлах плазмон является относительно долгоживущим возбуждением, и именно там он наиболее достоверно определяется через характеристические потери электронов.

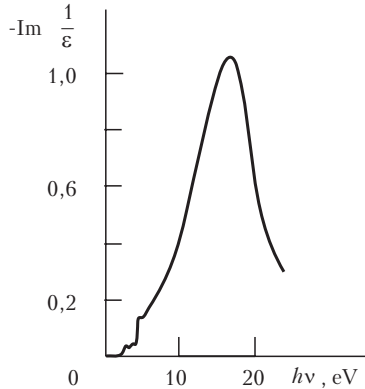


Рис. 6. Спектр характеристических потерь для Ge [40]

На примере плазмона можно наиболее просто оценить влияние поверхности на возбуждения в твердом теле, поскольку ее влияние на другие электронные возбуждения (образование таммовских состояний) оценить можно лишь на уровне микроскопической теории. Для плазменных волн в диэлектрике, ограниченном плоскостью и занимающем полупространство, уравнение для частоты продольных волн имеет вид

$$\tilde{\varepsilon}(\omega_{ps}) + 1 = 0. \quad (2.4)$$

В рамках приближений, для которых получено выражение (2.3), имеем (см., напр. [41])

$$\omega_{ps} = \sqrt{\frac{2\pi N e^2}{m}},$$

т. е. $\omega_{ps} = \omega_{pl}/\sqrt{2}$. В случае же, когда асимптотика (2.2) вблизи ω_{ps} не работает, необходимо пользоваться точным уравнением (2.4), а отношение частот поверхностных и объемных плазмонов отличается от $1/\sqrt{2}$.

Функция $\varepsilon(\omega)$ может обращаться в ноль не только в области энергий электронных переходов, но и в области фононных колебаний решетки. В первом приближении в фононной области диэлектрическая проницаемость ионных кристаллов имеет вид (см., например, [1], а также часть 3 данного пособия):

$$\varepsilon(\omega) = \varepsilon_\infty + \frac{(\varepsilon_0 - \varepsilon_\infty)\Omega_{\text{TO}}^2}{\Omega_{\text{TO}}^2 - \omega^2}. \quad (2.5)$$

При этом

$$\frac{\varepsilon_2(\omega)}{|\tilde{\varepsilon}(\omega)|^2} \approx \frac{\pi\Omega_{\text{LO}}}{2} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right) \left[\delta(\omega - \Omega_{\text{LO}}) - \delta(\omega + \Omega_{\text{LO}}) \right],$$

где ε_0 – статическая и ε_∞ – высокочастотная диэлектрические проницаемости, Ω_{LO} – частота продольного, а Ω_{TO} – поперечного оптического фонона, связанные между собой соотношением $\Omega_{\text{LO}}^2 = \Omega_{\text{TO}}^2 \frac{\varepsilon_0}{\varepsilon_\infty}$. Частота Ω_{TO} соответствует полюсу диэ-

лектрической проницаемости (2.5), а Ω_{LO} – нулю этой функции.

Функция потерь энергии электрона показывает в этом случае на испускание и поглощение продольных оптических фононов. Таким образом, быстрый электрон при прохождении через вещество теряет свою энергию на создание всего комплекса элементарных возбуждений, взаимодействующих с электроном через продольное электрическое поле: продольных оптических фононов, электронно-дырочных возбуждений и плазмонов. Проведенное выше рассмотрение было справедливо для электронов достаточно высоких энергий, таких, что испускание возбуждения с энергией в несколько электрон-вольт незначительно меняло импульс возбуждения. Для рассеяния более низкоэнергетичных электронов необходимо учитывать рождение не только длинноволновых возбуждений, но и коротковолновых, поэтому функция потерь уже не выражается через $\tilde{\varepsilon}(\omega)$, а лишь через $\tilde{\varepsilon}(\omega, \mathbf{q})$. Подробно все эти процессы рассмотрены в другой части пособия.

§ 3 Микроскопическое описание взаимодействия света с веществом

3.1 Диэлектрическая проницаемость

Взаимодействие света с веществом описывается обычно в полуклассическом приближении. Подробное описание этих процессов можно найти в ряде книг [42–45]. Здесь мы введем основные понятия, касающиеся этого взаимодействия. В полуклассическом приближении гамильтониан системы «электромагнитное поле плюс твердое тело» может быть разбит на сумму трех слагаемых:

$$\hat{H} = \hat{H}_0 + \hat{H}_R + \hat{H}_{\text{int}} , \quad (3.1)$$

где \hat{H}_0 — гамильтониан, описывающий твердое тело без излучения, \hat{H}_R — гамильтониан свободного электромагнитного поля, \hat{H}_{int} — член, описывающий взаимодействие света с веществом. Полуклассическое описание предусматривает, во-первых, квазиклассическое описание поля излучения. Во-вторых, член \hat{H}_{int} рассматривается лишь как возмущение. При таком описании опускаются эффекты образования поляритонов разного типа — состояний, представляющих собой смесь электронных или фоновых возбуждений с фотонами. Взаимодействие подобного типа на примере экситонного поляритона будет разобрано в следующей части.

Согласно тем же полуклассическим представлениям для вычисления \hat{H}_{int} поступают следующим образом. Электромагнитное поле рассматривается как внешняя система, поэтому член \hat{H}_R в (3.1) опускается. При помещении твердого тела в электромагнитное поле в гамильтониане, описывающем твердое тело, происходит замена операторов импульса всех частиц $\hat{\mathbf{p}}_i$ на операторы обобщенного импульса

$\hat{\mathcal{P}}_i = \hat{\mathbf{p}}_i + \frac{e_i}{c} \mathbf{A}(\mathbf{r}_i)$, где $\mathbf{A}(\mathbf{r}, t)$ – вектор-потенциал электромагнитного поля, e_i и \mathbf{r}_i – заряд и координаты i -й частицы. В нерелятивистском приближении обобщенные импульсы входят в гамильтониан в виде суммы кинетических энергий частиц:

$$\sum_i \frac{\hat{\mathbf{p}}_i^2}{2m_i} = \sum_i \frac{\hat{\mathcal{P}}_i^2}{2m_i} - \sum_i \frac{e_i}{m_i c} \mathbf{A}(\mathbf{r}_i) \hat{\mathcal{P}}_i + \sum_i \frac{e_i^2}{2m_i c^2} \mathbf{A}^2(\mathbf{r}_i). \quad (3.2)$$

При вычислении второго члена учтено, что для поперечной электромагнитной волны $\operatorname{div} \mathbf{A} = 0$, и поэтому $\mathbf{A}(\mathbf{r}_i)$ и $\hat{\mathcal{P}}_i$ коммутируют друг с другом. В случае относительно слабых интенсивностей электромагнитного излучения нелинейный по вектор-потенциалу член в (3.2) можно опустить. Его влияние становится существенным в основном при использовании мощного излучения лазеров. Таким образом, в полуклассическом приближении

$$\hat{\mathcal{H}}_{\text{int}} = -\sum_i \frac{e_i}{m_i c^2} \mathbf{A}(\mathbf{r}_i) \hat{\mathcal{P}}_i. \quad (3.3)$$

Выражение для диэлектрической проницаемости твердого тела можно получить многими способами. Мы выберем не очень строгий, но наглядный путь решения этой задачи. Более последовательно эти формулы будут рассмотрены в следующих разделах этого параграфа. Для начала рассмотрим поглощение на системе атомов, слабо взаимодействующих друг с другом.

Коэффициент поглощения $\alpha(\omega)$ можно выразить через сечение поглощения на одном атоме $\sigma(\omega) = \alpha(\omega)/N$, где N – число атомов в элементе объема. Для простоты рассмотрим плоскополяризованную падающую волну, распространяющуюся в направлении оси x (при этом $\hat{k}_x = (n + ik)\omega/c$). Вектор-потенциал такой электромагнитной волны равен

$$\mathbf{A}(\mathbf{r}, t) = A \mathbf{e} \cos(\omega t - \omega n x/c) \exp(-\omega k x/c)$$

где \mathbf{e} – вектор поляризации света, лежащий в плоскости yz . Поток энергии в электромагнитной волне описывается соотношением (1.16), а поток фотонов выражается величиной

$$\frac{S}{\hbar\omega} = \frac{A^2\omega n}{8\pi\hbar c} \exp\left(-\frac{2\kappa\omega x}{c}\right). \quad (3.4)$$

Гамильтониан взаимодействия вещества с такой электромагнитной волной можно записать в виде

$$\hat{\mathcal{H}}_{\text{int}} = \frac{e}{mc} A \sum_i e^{-\omega\kappa x_i/c} \cos(\omega t - \omega n x_i/c) \mathbf{e} \hat{\mathbf{p}}_i, \quad (3.5)$$

где суммирование проводится по всем электронам с номерами i . При этом предполагается, что взаимодействие происходит только с электронной подсистемой (e — величина заряда электрона, m — его масса). Непосредственное взаимодействие электромагнитной волны с ионной подсистемой не учитывается. Это возможно, если не рассматривать инфракрасную область спектра, в которой происходит прямое превращение фотона в поляризационный фонon.

Использование теории возмущений предполагает, что известны основное состояние и возбужденные состояния системы, описываемой гамильтонианом $\hat{\mathcal{H}}_0$. Каждое из этих состояний характеризуется набором квантовых чисел, который мы будем обозначать одной буквой, например, i для начального состояния $|i\rangle$ системы или f для конечного состояния $|f\rangle$. Ниже чаще всего будут использоваться дираковские обозначения векторов состояний (с использованием угловых скобок), но в некоторых случаях будут выписываться полные шредингеровские волновые функции, в том числе их явная зависимость от координат гильбертова пространства.

Среди квантовых чисел, которые характеризуют состояния системы с невозмущенным гамильтонианом $\hat{\mathcal{H}}_0$, часто удобно выделить энергию состояния E_α (для состояния $|\alpha\rangle$) и полный импульс \mathbf{P}_α . В случае стационарной однородной системы энергия и импульс, соответствующие собственным состояниям гамильтониана, сохраняются и являются хорошими квантовыми числами. Вообще говоря, проблема нахождения собственных состояний и соответствующих собственных значений энергии и полного импульса не имеет точного решения. Разработано большое число приближенных методов их нахождения. Здесь мы будем предполагать, что эта проблема уже решена, и известен полный набор состояний $|\alpha\rangle$ системы, соответствующие гамильтониану $\hat{\mathcal{H}}_0$, не возмущенному

внешним электромагнитным полем.

В полуклассическом приближении предполагается, что твердое тело находится в одном из состояний $|i\rangle$, и переход происходит в другое состояние $|f\rangle$ под действием внешнего электромагнитного поля. Вероятность перехода из состояния $|i\rangle$ в состояние $|f\rangle$ под действием возмущения (3.5) можно вычислить, используя «Золотое правило Ферми»:

$$\begin{aligned} w_{if} &= 2\pi\hbar^{-1} \left| \langle f | \hat{\mathcal{H}}_{\text{int}} | i \rangle \right|^2 \delta(\hbar\omega - E_f + E_i) = \\ &= \frac{\pi e^2}{2\hbar^2 m^2 c^2} A^2 \left| \langle f | \sum_j e^{i\omega n x_j / c - \omega k x_j / c} \hat{\mathbf{p}}_j \mathbf{e} | i \rangle \right|^2 \delta(\omega - \omega_{fi}), \end{aligned}$$

где $\omega_{fi} = (E_f - E_i)/\hbar$ — частота перехода из одного электронного состояния в другое. Чтобы получить парциальное сечение поглощения на одном атоме, необходимо w_{fi} поделить на поток фотонов (3.4):

$$\sigma_{if}(\omega) = \frac{4\pi^2 e^2}{\hbar m^2 c \omega n} \left| \langle f | \sum_j \hat{\mathbf{p}}_j \mathbf{e} | i \rangle \right|^2 \delta(\omega - \omega_{fi}), \quad (3.6)$$

Здесь предполагается, что работает дипольное приближение — длина волны излучения много больше характерных размеров волновых функций электронов. Поэтому в разложении экспоненты $e^{i\omega n x_j / c}$ оставлен только нулевой член. При этом матричный элемент $\langle f | \sum_j \hat{\mathbf{p}}_j | i \rangle$ считается отличным от нуля.

Если же симметрия начального и конечного состояний такова, что $\langle f | \sum_j \hat{\mathbf{p}}_j | i \rangle = 0$, необходимо учитывать следующие члены разложения экспоненты в ряд. Это приводит к процессам магнитно-дипольного и квадрупольного поглощения света.

Для преобразования выражения (3.6) можно воспользоваться известным коммутационным соотношением

$$\hat{\mathbf{p}}_j = \frac{im}{\hbar} [\hat{\mathcal{H}}, \mathbf{r}_j].$$

что приводит к равенству

$$\langle f | \sum_j \hat{\mathbf{p}}_j | i \rangle = im\omega_{fi} \langle f | \sum_j \mathbf{r}_j | i \rangle.$$

Для того чтобы перейти от сечения поглощения двухуровневой системы к коэффициенту поглощения электромагнитного излучения в твердом теле, необходимо учесть два обстоятельства. Во-первых, необходимо просуммировать (3.6) по всем возможным конечным состояниям в твердом теле, исключая занятые состояния (в соответствии с принципом Паули). Во-вторых, при ненулевой температуре твердое тело не находится в своем основном состоянии. При конечной температуре можно ввести функцию распределения по энергии $f(E)$, описывающую заселенность квантовых состояний. При этом коэффициент поглощения приобретает вид

$$\alpha(\omega) = \frac{4\pi^2 e^2 \omega N}{\hbar c n} \sum_{i,f} \left\langle \left| f \left| \sum_j \mathbf{r}_j \mathbf{e} \right| i \right\rangle \right|^2 \times \quad (3.7)$$

$$\times [1 - f(E_f)] f(E_i) \delta(\omega - \omega_{fi}),$$

а мнимая часть диэлектрической проницаемости становится равной

$$\varepsilon_2(\omega) = \frac{4\pi^2 e^2 N}{\hbar} \sum_{i,f} \left\langle \left| f \left| \sum_j \mathbf{r}_j \mathbf{e} \right| i \right\rangle \right|^2 \times \quad (3.8)$$

$$\times [1 - f(E_f)] f(E_i) \delta(\omega - \omega_{fi}),$$

Необходимо отметить, что в этом приближении правая часть уравнения (3.8) не содержит никаких оптических констант (в отличие от формулы (3.7), в знаменателе которой стоит показатель преломления $n(\omega)$). Это означает, что в приближении слабой связи конденсированной среды с электромагнитным полем переходы между различными состояниями вносят аддитивный вклад в мнимую часть диэлектрической проницаемости $\varepsilon_2(\omega)$. Это выделяет диэлектрическую проницаемость среди остальных оптических функций и делает ее наиболее удобной для теоретических расчетов.

Действительную часть диэлектрической проницаемости можно выразить через $\varepsilon_2(\omega)$ с помощью соотношений Крамера — Кронига. Тогда для комплексной $\tilde{\varepsilon}(\omega)$ получаем

$$\tilde{\varepsilon}(\omega) = 1 + \frac{8\pi e^2 N}{\hbar} \sum_{i,f} \left| \left\langle f \left| \sum_j \mathbf{r}_j \mathbf{e} \right| i \right\rangle \right|^2 \times \quad (3.9)$$

$$\times \frac{\omega_{fi}}{\omega_{fi}^2 - (\omega + is)^2} \Big|_{s \rightarrow +0} [1 - f(E_f)] f(E_i),$$

В этом выражении введена бесконечно малая положительная: величина s , определяющая правильный знак ε_2 . В реальных случаях, поскольку возбужденные состояния в твердом теле не имеют бесконечного времени жизни, роль s играет величина $2\pi/\tau_f$, где τ_f — время жизни данного возбужденного состояния. Более подробно описание возбужденных состояний с конечным временем жизни будет дано при исследовании формы линии экситонного поглощения.

Выражение (3.9) дает возможность проанализировать, какие переходы вносят вклад в действительную и мнимую части диэлектрической проницаемости. В формуле (3.8) для мнимой части диэлектрической проницаемости стоит дельта-функция, соответствующая закону сохранения энергии для реальных переходов между различными состояниями. Поэтому в ε_2 в этом приближении вносят вклад только переходы между такими состояниями, разница энергий между которыми находится в резонансе с падающим светом. С другой стороны, в действительную часть диэлектрической проницаемости ε_1 вносят вклад все возможные переходы, в том числе и те, для которых не выполняется закон сохранения энергии при переходе. Вклад таких состояний пропорционален

$$\Delta t \approx 2\pi\hbar / |\hbar\omega - E_f + E_i| = 2\pi / |\omega - \omega_{fi}|$$

— времени, в течение которого может нарушаться закон сохранения энергии (в соответствии с соотношением неопределенности между наблюдаемой энергией ΔE и временем измерения Δt , связанных между собой соотношением $\Delta E \Delta t \approx 2\pi\hbar$). Такая интерпретация дает возможность говорить, что действительная часть диэлектрической проницаемости ε_1 соответствует виртуальным переходам. В частности, диэлектрическая проницаемость в области прозрачности, которая является чисто действительной величиной, определяется в основном виртуальными переходами с энергией в области

фундаментального поглощения.

При выводе формул (3.6) – (3.9) предполагалось, что волновые функции конечного и начального состояния нормированы на единицу для каждого из атомов, поэтому в качестве множителя выступает N – число атомов в единице объёма. Если пользоваться волновыми функциями твердого тела, нормированными на единичный объем, множитель N пропадает. Первый способ удобен для описания вклада в диэлектрическую проницаемость локализованных электронных состояний, таких, как электронные состояния дефектов и примесей в твердом теле. Второй способ необходим при описании собственных электронных состояний, например состояний непрерывного спектра и т. д.

Помимо различия в нормировке волновых функций вклад локализованных состояний отличается от вклада состояний, принадлежащих всему кристаллу, еще одним множителем – так называемой поправкой на эффективное поле. Эта поправка обычно записывается как $(E_{\text{eff}}/E_0)^2$. Происхождение ее следует из того факта, что эффективное поле E_{eff} , которое индуцирует переходы, содержит вклад, пропорциональный поляризации, и не равно приложенному полю E_0 . Простая модель Лоренц – Лоренца, по которой вычисляется электрическое поле в сферической полости, находящейся в диэлектрической среде с диэлектрической проницаемостью ϵ_∞ , дает $E_{\text{eff}}/E_0 = (\epsilon_\infty + 2)/3$. При этом в качестве диэлектрической проницаемости надо брать высокочастотную проницаемость ϵ_∞ , поскольку ионы практически не успевают за изменением электрического поля. В отличие от статической диэлектрической проницаемости ϵ_0 , в которую вносят вклады как электронные переходы, так и колебания ионов, ϵ_∞ соответствует области прозрачности кристалла и связана только с электронными переходами. Символ ∞ означает, что соответствующие частоты много больше характерных частот движения ионов в твердом теле. Эти обозначения, уже встречавшиеся в предыдущем разделе, будут широко использоваться при рассмотрении экситонных и фононных эффектов.

В действительности корректно учесть поправку на эффективное поле достаточно трудно, поскольку в области электронных переходов необходимо учитывать частотную дисперсию $\epsilon(\omega)$, несферичность полости, в которой находится дефект, и т. д. Разные методы дают довольно сильно

различающиеся результаты. Экспериментальная проверка в ряде случаев приводит к заключению, что формула Лоренц—Лоренца переоценивает поправку на эффективное поле (см., напр., [5]). Кроме того, эта поправка, возможно, оказывается разной для различных состояний дефекта, в частности из-за неодинаковой протяженности состояний. Здесь мы ограничимся сделанными комментариями.

Сделанные выше замечания указывают на важность пространственной дисперсии, то есть различного отклика диэлектрической среды под воздействием электромагнитного поля с различными пространственными частотами. В действительности, поправка на эффективное поле связана с попыткой найти такую амплитуду длинноволнового электромагнитного поля, которая по своему действию была бы эквивалентна реальной электромагнитной волне в твердом теле. Такая волна содержит большое количество коротковолновых компонент, быстро осциллирующих на межатомных расстояниях (более подробно см. [10]). Корректный учет этого явления возможен только при расчете диэлектрической проницаемости $\epsilon(\omega, \mathbf{k})$ для всех значений волнового вектора \mathbf{k} . При учете зависимости от \mathbf{k} свойства диэлектрической проницаемости существенно изменяются. В частности, в случае сильной пространственной дисперсии величина $\epsilon(\omega, \mathbf{k})$ не удовлетворяет соотношениям Крамерса-Кронига (эти соотношения всегда выполняются лишь для функции $\epsilon^{-1}(\omega, \mathbf{k})$) и т.д. Более подробно эти вопросы освещены в работах [46, 47].

3.2 Поляризационное приближение для взаимодействия электромагнитных полей с твердым телом

В предыдущем разделе, исходя из простых предположений, было получено выражение (3.9) для диэлектрической проницаемости $\epsilon(\omega)$. В данном разделе будет обобщен вывод соответствующих уравнений, при этом основное внимание будет уделяться другой стороне взаимодействия электромагнитного поля с веществом. В основном это рассмотрение будет основано на работах Ю.Л.Климонтовича [48]).

Как и в разделе 3.1, будем предполагать, что система описывается гамилтонианом (3.1), и что собственные значения и собственные векторы \hat{H}_0 известны. Матрица плотности системы без учета взаимодействия будет обозначаться через \mathcal{R}_{ij} .

Уравнение для матрицы плотности имеет обычный вид [49]:

$$i\hbar \frac{\partial \mathcal{R}_{nm}}{\partial t} = \sum_k (\hat{\mathcal{H}}_{nk} \mathcal{R}_{km} - \mathcal{R}_{nk} \hat{\mathcal{H}}_{km}) .$$

Будем предполагать, что состояния n , используемые в качестве базисных, соответствуют многочастичным собственным значениям гамильтониана $\hat{\mathcal{H}}_0$. Тогда уравнение для матрицы плотности системы, описываемой гамильтонианом с учетом взаимодействия, имеет форму

$$\begin{aligned} \left(\frac{\partial}{\partial t} + i\omega_{nm} \right) \mathcal{R}_{nm} = \\ = -\frac{i}{\hbar} \sum_k \left((\hat{\mathcal{H}}_{\text{int}})_{nk} \mathcal{R}_{km} - \mathcal{R}_{nk} (\hat{\mathcal{H}}_{\text{int}})_{km} \right) . \end{aligned} \quad (3.10)$$

В равновесном состоянии матрица плотности системы диагональна. Таким образом, для состояний, достаточно близких к равновесию, можно записать

$$\mathcal{R}_{nm} = f_n \delta_{nm} + \delta f_{nm} . \quad (3.11)$$

Существенно более аккуратное описание эволюции матрицы плотности, в том числе и оценка точности примененных приближений, может быть получена другими методами, например, с использованием техники проекционных операторов (см., например, [50–52]). Приближение (3.11) соответствует рассмотрению системы на временах, больших так называемых "времен поперечной релаксации", в течение которых пропадают недиагональные элементы матрицы плотности.

В состоянии равновесия среднее значение $\langle f_n \rangle = f_n^0$ удовлетворяет распределению Гиббса

$$f_n^0 = \frac{1}{Z} e^{-E_n/k_B T} , \quad Z = \sum_n e^{-E_n/k_B T} , \quad (3.12)$$

где k_B — постоянная Больцмана, а T — абсолютная температура. Отметим, что в данном разделе f_n описывает вероятность нахождения всей системы в многочастичном состоянии n и не имеет никакого отношения к числам заполнения одноэлектронных состояний, появляющихся в приближении Хартри–Фока, которые будут обозначаться как f_α . Тем самым при

нулевой температуре ($T = 0$) $f_n = \delta_{n0}$, что означает нахождение всей системы в основном состоянии.

Выражение для оператора взаимодействия (3.3) может быть записано через оператор тока

$$\hat{\mathbf{j}}(\mathbf{r}) = \sum_i \frac{e_i}{m_i} \hat{\mathcal{P}}_i \delta(\mathbf{r} - \mathbf{r}_i) ,$$

где суммирование проводится по всем частицам с номерами i . Тогда

$$\hat{\mathcal{H}}_{\text{int}} = -\int \mathbf{A}(\mathbf{r}, t) \hat{\mathbf{j}}(\mathbf{r}) d^3 r . \quad (3.13)$$

Оператор тока $\hat{\mathbf{j}}$ связан с вектором поляризации $\mathbf{P}(\mathbf{r}, t)$ соотношением $\hat{\mathbf{j}} = \frac{\partial \mathbf{P}(\mathbf{r}, t)}{\partial t}$. Подставляя это равенство в (3.13) и выделяя полную производную по времени, получаем

$$\hat{\mathcal{H}}_{\text{int}} = -\int \mathbf{E}(\mathbf{r}, t) \mathbf{P}(\mathbf{r}, t) d^3 r + \frac{\partial}{\partial t} (\dots) . \quad (3.14)$$

В дальнейшем нам понадобятся фурье-компоненты $\hat{\mathcal{H}}_{\text{int}}$ по времени. В этом случае полную производную можно опустить.

Вектор поляризации связан с матрицей плотности системы равенством

$$\mathbf{P}(\mathbf{r}, t) = \sum_{nm} \mathbf{D}_{mn}(\mathbf{r}) \mathcal{R}_{nm}(t) = \sum_{nm} \mathbf{D}_{mn}(\mathbf{r}) \delta f_{nm}(t) , \quad (3.15)$$

где \mathbf{D}_{nm} — матричный элемент оператора дипольного момента

$$\hat{\mathbf{D}}(\mathbf{r}) = \sum_i e_i \mathbf{r}_i \delta(\mathbf{r} - \mathbf{r}_i) , \quad \mathbf{D}_{nm} = \langle n | \hat{\mathbf{D}}(\mathbf{r}) | m \rangle . \quad (3.16)$$

Последнее равенство в (3.15) связано с тем, что в равновесном состоянии вектор поляризации всей системы обращается в ноль (сегнетоэлектрические системы здесь не рассматриваются).

Флуктуационные добавки к матрице плотности разделяются на сумму двух слагаемых: индуцированных полем $\delta f_{nm}^{\text{ind}}$ и флуктуаций источника $\delta f_{nm}^{\text{src}}$. Флуктуации источника имеют фундаментальную природу и связаны с атомарной структурой вещества. Тем самым и поляризация \mathbf{P} в свою очередь разбивается на две части: поляризацию \mathbf{P}^{ind} , индуцированную

электромагнитным полем, и поляризацию источника \mathbf{P}^{src} .

Важную роль играют величины, усредненные по ансамблю систем. Такое усреднение будем обозначать угловыми скобками $\langle \dots \rangle$. Чтобы получить выражение для корреляционных функций флуктуаций источника $\langle \delta f_{nm}^{\text{src}}(t) \delta f_{n'm'}^{\text{src}*}(t') \rangle$, можно рассмотреть двухвременную матрицу плотности системы $\mathcal{R}_{nm,n'm'}(t, t')$. При совпадающих временах эта матрица выражается через $\mathcal{R}_{nm}(t)$:

$$\mathcal{R}_{nm,n'm'}(t, t) = \frac{1}{2} [\delta_{nn'} \mathcal{R}_{mm'}(t) + \delta_{mm'} \mathcal{R}_{nn'}(t)] .$$

Корреляционная функция флуктуаций источника определяется как разница между двухвременной матрицей плотности и произведением двух одновременных матриц плотности:

$$\langle \delta f_{nm}^{\text{src}}(t) \delta f_{n'm'}^{\text{src}*}(t') \rangle = \mathcal{R}_{nm,n'm'}(t, t') - \mathcal{R}_{nm}(t) \mathcal{R}_{n'm'}(t') .$$

Отсюда для совпадающих времен получаем

$$\begin{aligned} \langle \delta f_{nm}^{\text{src}}(t) \delta f_{n'm'}^{\text{src}*}(t) \rangle &= \\ &= \frac{1}{2} [\delta_{nn'} \mathcal{R}_{mm'}(t) + \delta_{mm'} \mathcal{R}_{nn'}(t)] - \mathcal{R}_{nm}(t) \mathcal{R}_{n'm'}(t) . \end{aligned}$$

Если рассматриваемое состояние неравновесно, но достаточно близко от состояния равновесия, то

$$\begin{aligned} \langle \delta f_{nm}^{\text{src}}(t) \delta f_{n'm'}^{\text{src}*}(t) \rangle &= \\ &= \frac{1}{2} \delta_{nn'} \delta_{mm'} [f_m(t) + f_n(t)] - \delta_{nm} \delta_{n'm'} f_n(t) f_n'(t) . \end{aligned} \quad (3.17)$$

Эта формула аналогична тому, что при пуассоновском распределении средний квадрат числа событий $\langle (\Delta N)^2 \rangle = \langle N^2 \rangle - \langle N \rangle^2$ равен $\langle N \rangle$.

В формуле (3.17) предполагается, что временная зависимость функций $f_n(t)$ относительно медленная (по сравнению с временами порядка $2\pi\omega_{nm}^{-1}$).

Флуктуационные источники $\delta f_{nm}^{\text{src}}$ имеются и в системе без взаимодействия ($\hat{\mathcal{H}}_{\text{int}} = 0$), и поэтому коррелятор удовлетворяет уравнению

$$\left(\frac{\partial}{\partial t} + i\omega_{nm} + s \right) \langle \delta f_{nm}^{\text{src}}(t) \delta f_{n'm'}^{\text{src}*}(t') \rangle = 0$$

с начальным условием (3.17) ($s > 0$ – малая добавка, обеспечивающая ослабление корреляций на больших временах). Предполагая стационарность системы, можно перейти к фурье-компонентам по времени, и для $\omega \neq 0$ получаем

$$(\delta f_{nm}^{\text{src}} \delta f_{n'm'}^{\text{src}})_{\omega} = \pi \delta(\omega - \omega_{nm}) \delta_{nn'} \delta_{mm'} (f_m + f_n), \quad (3.18)$$

где $(\dots)_{\omega}$ обозначает спектральную плотность, определяемую в стационарном случае соотношением

$$\langle x_{\omega} y_{\omega'} \rangle = \langle x_{\omega} y_{-\omega'}^* \rangle = 2\pi (xy)_{\omega} \delta(\omega + \omega'), \quad (3.19)$$

и связанную с корреляционной функцией $\langle x(t + \tau)y(t) \rangle$ (в стационарном случае выполняется соотношение $\langle x(t + \tau)y(t) \rangle = \langle x(\tau)y(0) \rangle$, то есть корреляционная функция зависит только от τ и не зависит от времени t) Фурье-преобразованием

$$(xy)_{\omega} = \int_{-\infty}^{\infty} \langle x(\tau)y(0) \rangle e^{i\omega\tau} d\tau. \quad (3.20)$$

Используя обратное преобразование, можно выразить корреляционную функцию через спектральную плотность:

$$\begin{aligned} \langle x(t)y(0) \rangle &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} (xy)_{\omega} e^{-i\omega t} d\omega = \frac{1}{2\pi} \int_0^{\infty} 2(xy)_{\omega} e^{-i\omega t} d\omega. \end{aligned} \quad (3.21)$$

Здесь отдельно выписано интегрирование только по положительным частотам, что часто используется при описании спектральных характеристик. Необходимо напомнить, что

$$(xy)_{\omega} = (yx)_{-\omega} = (yx)_{\omega}^*. \quad (3.22)$$

Если рассматривается нестационарный случай, то f_n медленно меняются со временем, и формулы (3.18) – (3.22) выполняются лишь приближенно, с точностью до отношения быстрых характерных времен $2\pi\omega_{nm}^{-1}$ ко времени релаксации. Малость этого параметра связана с предположением малости выделенного гамильтониана взаимодействия $\hat{\mathcal{H}}_{\text{int}}$. В таком случае корреляционные функции типа $\langle x(t + \tau)y(t) \rangle$ слабо зависят от времени t и существенно – от времени τ . Переходя к

Фурье-компонентам по τ , получаем спектральные характеристики, медленно меняющиеся со временем t . Такое двойное представление в переменных время-частота возможно в случае существенной разницы характерных времен системы. Медленная зависимость от t будет выделяться отделением этого аргумента функций точкой с запятой от остальных переменных. Кроме того, предполагается, что возмущение системы внешним полем является слабым, и возбужденные состояния гамильтониана $\hat{\mathcal{H}}_0$ не изменяются. В частности, со временем не меняются состояния системы $|n\rangle$, их энергии E_n и матричные элементы \mathbf{D}_{nm} .

Если перейти к Фурье-компонентам и предположить малость матричных элементов $\hat{\mathcal{H}}_{\text{int}}$, то из (3.10) можно получить уравнение для индуцированной внешним полем добавки к функции распределения:

$$\begin{aligned} (-i\omega + s + i\omega_{nm}) \delta f_{nm}^{\text{ind}}(\omega; t) = \\ = \frac{i}{\hbar} (f_m(t) - f_n(t)) \int \mathbf{D}_{nm}(\mathbf{r}) \mathbf{E}(\mathbf{r}, \omega; t) d^3 r, \end{aligned}$$

где вместо \mathcal{R}_{nm} используется ее диагональное приближение, и для сходимости интегралов Фурье в уравнения добавлен малый релаксационный член, пропорциональный $s \rightarrow 0$. Отсюда полная добавка к функции распределения имеет вид

$$\begin{aligned} \delta f_{nm}(\omega; t) = \delta f_{nm}^{\text{src}}(\omega; t) + \\ = \frac{f_n(t) - f_m(t)}{\hbar(\omega + is - \omega_{nm})} \int \mathbf{D}_{nm}(\mathbf{r}) \mathbf{E}(\mathbf{r}, \omega; t) d^3 r. \end{aligned} \quad (3.23)$$

Ниже в уравнениях медленная зависимость от времени t будет часто опускаться.

Учитывая (3.15), получаем соотношение между поляризацией и электрическим полем:

$$\begin{aligned} P_i(\mathbf{r}, \omega; t) = \\ = \int \tilde{\alpha}_{ij}(\mathbf{r}, \mathbf{r}', \omega; t) E_j(\mathbf{r}', \omega; t) d^3 r' + P_i^{\text{src}}(\mathbf{r}, \omega; t). \end{aligned} \quad (3.24)$$

Здесь введен тензор поляризуемости

$$\tilde{\alpha}_{ij}(\mathbf{r}, \mathbf{r}', \omega; t) = \sum_{nm} (D_i(\mathbf{r}))_{nm} (D_j(\mathbf{r}'))_{nm} \frac{f_n(t) - f_m(t)}{\hbar(\omega + is - \omega_{nm})}. \quad (3.25)$$

Поляризационный источник \mathbf{P}^{src} связан с источником $\delta f_{nm}^{\text{src}}$ формулой, аналогичной (3.15). Учитывая соотношение (1.3), можно получить тензор диэлектрической проницаемости через тензор поляризуемости: $\tilde{\epsilon}_{ij} = 1 + 4\pi\tilde{\alpha}_{ij}$.

В формуле (3.24) можно перейти к Фурье-компонентам по волновому вектору \mathbf{k} :

$$P_i(\omega, \mathbf{k}) = \frac{1}{(2\pi)^3} \int \alpha_{ij}(\mathbf{k}, \mathbf{k}', \omega) E_j(\omega, \mathbf{k}') d^3k' + P_i^{\text{src}}(\omega, \mathbf{k}),$$

где

$$\begin{aligned} \tilde{\alpha}_{ij}(\mathbf{k}, \mathbf{k}', \omega) = \\ = \sum_{nm} (D_i(-\mathbf{k}))_{nm}^* (D_j(-\mathbf{k}'))_{nm} \frac{f_n - f_m}{\hbar(\omega + is - \omega_{nm})}. \end{aligned} \quad (3.26)$$

В пространственно-однородной системе тензор поляризуемости (3.25) зависит только от разности $\mathbf{r} - \mathbf{r}'$, в соответствии с чем

$$\tilde{\alpha}_{ij}(\mathbf{k}, \mathbf{k}', \omega) = (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}') \tilde{\alpha}_{ij}(\omega, \mathbf{k}). \quad (3.27)$$

Если, кроме того, система изотропна, $\tilde{\alpha}_{ij}(\omega, \mathbf{k})$ сводится к сумме двух тензоров, один из которых пропорционален δ_{ij} , а другой $-k_i k_j / \mathbf{k}^2$ (в изотропной системе невозможно построить никаких других тензоров):

$$\tilde{\alpha}_{ij} = \tilde{\alpha}^{\parallel} k_i k_j / \mathbf{k}^2 + \tilde{\alpha}^{\perp} (\delta_{ij} - k_i k_j / \mathbf{k}^2), \quad (3.28)$$

где введены два скаляра: продольная $\tilde{\alpha}^{\parallel}$ и поперечная $\tilde{\alpha}^{\perp}$ поляризуемости. В таком случае, если $\tilde{\alpha}^{\parallel} \neq \tilde{\alpha}^{\perp}$, то зависимость $\tilde{\alpha}_{ij}$ от \mathbf{k} становится неаналитической в пределе $\mathbf{k} \rightarrow 0$. Умножая (3.28) на $k_i k_j$ и производя суммирование по i и j , можно получить выражение для продольной поляризуемости

$$\tilde{\alpha}^{\parallel}(\omega, \mathbf{k}) = \frac{1}{\hbar k^2} \sum_{nm} \left(|\tilde{\rho}_{nm}|^2 \right)_{-\mathbf{k}} \frac{f_n - f_m}{\omega + is - \omega_{nm}}, \quad (3.29)$$

где было использовано уравнение непрерывности (1.4), из которого для изотропной системы из него следует соотношение

$$\left(|\tilde{\rho}_{nm}|^2 \right)_{-\mathbf{k}} = \left(|\mathbf{k} \mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} = \frac{k^2}{3} \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}}.$$

Далее в большинстве случаев мы будем предполагать, что $\tilde{\alpha}^{\parallel} = \tilde{\alpha}^{\perp}$. В таком случае $\tilde{\alpha}_{ij} = \tilde{\alpha} \delta_{ij}$, откуда

$$\tilde{\alpha}(\omega, \mathbf{k}) = \frac{1}{3\hbar} \sum_{nm} \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} \frac{f_n - f_m}{\omega + is - \omega_{nm}}. \quad (3.30)$$

Поскольку в соответствии с соотношениями (3.27) дельта-функция $\delta(\mathbf{k} - \mathbf{k}')$ уже выделена из поляризуемости, обозначение $\left(|\cdot|^2 \right)_{-\mathbf{k}}$ в уравнениях (3.29) и (3.30) имеют тот же смысл, что и спектральная плотность $(x^2)_{\omega}$ в формуле (3.19), то есть являются коэффициентами перед соответствующей дельта-функцией, например,

$$\left(\tilde{\rho}(-\mathbf{k}) \right)_{nm}^* \left(\tilde{\rho}(-\mathbf{k}') \right)_{nm} = (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}') \left(|\tilde{\rho}_{nm}|^2 \right)_{-\mathbf{k}}.$$

В таких обозначениях диэлектрическая проницаемость принимает вид

$$\tilde{\epsilon}(\omega, \mathbf{k}) = 1 + \frac{4\pi}{3\hbar} \sum_{nm} \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} \frac{f_n - f_m}{\omega + is - \omega_{nm}}. \quad (3.31)$$

Ее мнимая часть равна

$$\epsilon_2(\omega, \mathbf{k}) = 1 + \frac{4\pi^2}{3\hbar} \sum_{nm} \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} (f_m - f_n) \delta(\omega - \omega_{nm}). \quad (3.32)$$

Меняя индексы, по которым происходит суммирование, $n \leftrightarrow m$, и используя очевидное соотношение $\omega_{nm} = -\omega_{mn}$, легко показать нечетность мнимой части $\tilde{\epsilon}$ и четность действительной: $\tilde{\epsilon}(\omega, \mathbf{k}) = \tilde{\epsilon}^*(-\omega, -\mathbf{k})$. Обобщение этого соотношения на анизотропный случай имеет вид:

$$\tilde{\epsilon}_{ij}(\omega, \mathbf{k}) = \tilde{\epsilon}_{ji}^*(-\omega, -\mathbf{k}).$$

В неравновесном состоянии система в общем случае неоднородна и неизотропна. В таком случае необходимо использовать общее выражение

$$\begin{aligned} \tilde{\epsilon}_{ij}(\mathbf{r}, \mathbf{r}', \omega, t) &= \delta(\mathbf{r} - \mathbf{r}') + \\ &+ 4\pi \sum_{nm} \left(D_i(\mathbf{r}) \right)_{nm} \left(D_j(\mathbf{r}') \right)_{nm} \frac{f_n(t) - f_m(t)}{\hbar(\omega + is - \omega_{nm})} \end{aligned} \quad (3.33)$$

в координатном представлении и формулу, аналогичную (3.26) в представлении волновых векторов. Отличие неравновесной диэлектрической проницаемости от равновесной описывает целый ряд эффектов, в частности, наведенное поглощение.

Наличие k^2 в знаменателе уравнения (3.29) связано с экранированием электромагнитного поля, если

$$\lim_{\mathbf{k} \rightarrow 0} \left(|\tilde{\rho}_{nm}|^2 \right)_{-\mathbf{k}} \neq 0. \quad (3.34)$$

В таком случае в длинноволновом пределе знаменатель первого члена в правой части уравнения (1.5) не зависит от \mathbf{k} , и электрическое поле точечного заряда падает быстрее любой степени расстояния (в большинстве случаев экспоненциально). Можно показать, что для локализованных состояний предел (3.34) равен нулю. Тем самым экранирование кулоновского поля напрямую связано с наличием делокализованных возбуждений, например, свободных электронов.

3.3 Флуктуационно-диссипационная теорема

Для дальнейшего обсуждения необходимо выразить электрическое поле через флуктуации матрицы плотности. Используя приведенную выше формулу (3.18) для коррелятора $\delta f_{nm}^{\text{src}}$, можно получить

$$(\delta \mathbf{P} \delta \mathbf{P})_{\omega, \mathbf{k}}^{\text{src}} = \pi \sum_{nm} \delta(\omega - \omega_{nm}) \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} (f_n + f_m). \quad (3.35)$$

С помощью формулы (1.7) получаем спектральную плотность флуктуаций продольных компонент поля $\mathbf{E}^{\parallel}(\omega, \mathbf{k}) \parallel \mathbf{k}$:

$$\begin{aligned} (\delta \mathbf{E}^{\parallel} \delta \mathbf{E}^{\parallel})_{\omega, \mathbf{k}} &= \\ &= \frac{16\pi^3}{k^2} \sum_{nm} \frac{\left(|\tilde{\rho}_{nm}|^2 \right)_{-\mathbf{k}}}{|\tilde{\epsilon}(\omega, \mathbf{k})|^2} (f_n + f_m) \delta(\omega - \omega_{nm}) \end{aligned} \quad (3.36)$$

и спектральную плотность флуктуаций поперечных компонент $\mathbf{E}^{\perp}(\omega, \mathbf{k}) \perp \mathbf{k}$ (таких компонент две, и они взаимно перпендикулярны):

$$\begin{aligned}
& (\delta \mathbf{E}^\perp \delta \mathbf{E}^\perp)_{\omega, \mathbf{k}}^{\text{src}} = \\
& = \frac{32\pi^3}{3} \sum_{nm} \frac{(|\mathbf{D}_{nm}|^2)_{-\mathbf{k}}}{|\tilde{\epsilon}(\omega, \mathbf{k}) - c^2 \mathbf{k}^2 / \omega^2|^2} (f_n + f_m) \delta(\omega - \omega_{nm}). \quad (3.37)
\end{aligned}$$

Эти спектральные плотности являются четными функциями частоты ω . Формулы (3.36) и (3.37) описывают флуктуации не только в равновесной среде, но также и в среде, находящейся в неравновесном состоянии (но достаточно близко от равновесного состояния, что более подробно рассмотрено в [48]). Эти формулы справедливы, если неравновесное состояние однородно и изотропно.

Формулы (3.35) – (3.37) могут быть обобщены на неоднородный случай в приближении отсутствия пространственной дисперсии, то есть когда $\tilde{\epsilon}(\omega, \mathbf{k}) = \tilde{\epsilon}(\omega)$. В таком случае координатная зависимость коррелятора компонент электрического поля имеют вид

$$\begin{aligned}
& \langle E_i(\omega, \mathbf{r}; t) E_j^*(\omega, \mathbf{r}'; t) \rangle_\omega = \\
& = \pi \int d^3 r_1 \int d^3 r_2 K_{ik}(\mathbf{r} - \mathbf{r}_1) K_{jl}^*(\mathbf{r}' - \mathbf{r}_2) \times \\
& \times \sum_{nm} \delta(\omega - \omega_{nm}) (\mathbf{D}_{nm}(\mathbf{r}_1))_k (\mathbf{D}_{mn}^*(\mathbf{r}_2))_l (f_m(t) + f_n(t)), \quad (3.38)
\end{aligned}$$

где ядро $K_{ij}(\mathbf{r})$ определено соотношением (1.12) и включает в себя квазистационарное поле диполя и поперечное поле излучения этого диполя.

В состоянии равновесия f_n^0 описывается распределением Гиббса (3.12), откуда легко получить

$$(f_m^0 + f_n^0) = (f_m^0 - f_n^0) \operatorname{cth} \frac{\hbar \omega_{nm}}{2k_B T}. \quad (3.39)$$

Подстановка этого выражения в (3.36) и (3.37) приводит к соотношениям Каллена – Велтона (см., например, [53]):

$$\begin{aligned}
(\delta E^{\parallel} \delta E^{\parallel})_{\omega, \mathbf{k}} &= \frac{4\pi\hbar \varepsilon_2(\omega, \mathbf{k})}{|\tilde{\varepsilon}(\omega, \mathbf{k})|^2} \operatorname{cth} \frac{\hbar\omega}{2k_B T} = \\
&= 4\pi\hbar \operatorname{Im} \left(-\frac{1}{\tilde{\varepsilon}(\omega, \mathbf{k})} \right) \operatorname{cth} \frac{\hbar\omega}{2k_B T},
\end{aligned} \tag{3.40}$$

$$\begin{aligned}
(\delta E^{\perp} \delta E^{\perp})_{\omega, \mathbf{k}} &= \frac{8\pi\hbar \varepsilon_2(\omega, \mathbf{k})}{|\tilde{\varepsilon}(\omega, \mathbf{k}) - c^2 \mathbf{k}^2 / \omega^2|^2} \operatorname{cth} \frac{\hbar\omega}{2k_B T} = \\
&= 8\pi\hbar \operatorname{Im} \left(-\frac{1}{\tilde{\varepsilon}(\omega, \mathbf{k}) - c^2 \mathbf{k}^2 / \omega^2} \right) \operatorname{cth} \frac{\hbar\omega}{2k_B T}.
\end{aligned} \tag{3.41}$$

Эти соотношения выражают флуктуационно-диссипационную теорему для электрического поля в среде.

В соотношения (3.40), (3.41) входит гиперболический котангенс, который можно записать в виде

$$\operatorname{cth} \frac{\hbar\omega}{2k_B T} = 2 \left[\frac{1}{2} + \frac{1}{\exp(\hbar\omega/k_B T) - 1} \right]. \tag{3.42}$$

Первый член в этой сумме соответствует нулевым колебаниям поля, а второй член соответствует распределению Бозе для фононов.

Флуктуации продольного поля отличны от нуля только для таких частот, для которых мнимая часть диэлектрической проницаемости не обращается в ноль. Кроме того, флуктуации резко возрастают, когда $\tilde{\varepsilon}$ приближается к нулю. Это соответствует флуктуационному возбуждению плазменных мод (см. выражение (1.13) и параграф 2). Флуктуации поперечного поля не равны нулю даже при $\varepsilon_2 = 0$, то есть в области прозрачности вещества. Вычисляя предел, соответствующий обращению в ноль знаменателя в соответствии с (1.14), получаем

$$\begin{aligned}
&\operatorname{Im} \left[-\frac{1}{\varepsilon_1(\omega) + i0 \cdot \operatorname{sign} \omega - c^2 \mathbf{k}^2 / \omega^2} \right] = \\
&= \frac{\pi\omega}{2cn(\omega)} \left[\delta(|\mathbf{k}| - \omega n(\omega)/c) + \delta(|\mathbf{k}| + \omega n(\omega)/c) \right],
\end{aligned} \tag{3.43}$$

где $n(\omega) = \sqrt{\varepsilon_1(\omega, 0)}$ — показатель преломления (1.10) в

области прозрачности. Интегрируя по всем волновым векторам, можно получить

$$\begin{aligned} 2(\mathbf{E}^\perp \mathbf{E}^\perp)_\omega &= \frac{1}{(2\pi)^3} \int d^3k \, 2(\mathbf{E}^\perp \mathbf{E}^\perp)_{\omega, \mathbf{k}} = \\ &= \frac{8\hbar\omega^3 n}{c^3} \left[\frac{1}{2} + \frac{1}{\exp(\hbar\omega/k_B T) - 1} \right]. \end{aligned}$$

Коэффициент 2 появляется из-за того, что интегрирование проводится только по положительным частотам. Учитывая, что на интервал $d\omega$ приходится энергия (см. [54])

$$\frac{1}{8\pi} \left[2(\mathbf{E}^\perp \mathbf{E}^\perp)_\omega \frac{d(\omega\varepsilon)}{d\omega} + 2(\mathbf{H}^\perp \mathbf{H}^\perp)_\omega \right] \frac{d\omega}{2\pi},$$

и, кроме того, в поперечной волне $|\mathbf{H}| = n|\mathbf{E}|$, получаем, что энергия \mathcal{U} электромагнитного поля в равновесном состоянии при температуре T в единице объема в частотном интервале $d\omega$ равна

$$\frac{d\mathcal{U}}{d\omega} = \left[\frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} \right] \frac{\omega^2 n^2}{\pi^2 c^3} \frac{d(n\omega)}{d\omega}.$$

Это выражение соответствует энергии черного излучения в среде с показателем преломления $n(\omega)$ с учетом нулевых колебаний (см. [55]).

В неравновесном случае флуктуации поля (3.36), (3.37) отличаются от равновесных (3.40), (3.41) из-за отличия функции распределения f_n от равновесной f_n^0 . Это отличие входит в формулы (3.36) и (3.37) не только явным образом, но также и через отличие диэлектрической проницаемости (3.31) от равновесной. Люминесценция твердых тел как раз и является превышением флуктуаций поперечного поля над уровнем тепловых флуктуаций.

3.4 Кинетическое уравнение в поляризационном приближении

Из уравнений (3.10) можно получить кинетическое уравнение для функции f_n . После усреднения правых частей уравнений (3.10) для диагональных компонент матрицы плотности члены, содержащие f_n , сокращаются, и остаются только

средние значения типа $\langle \hat{\mathcal{H}}_{\text{int}} \delta f_{nm} \rangle$. Используя выражение для гамильтониана взаимодействия (3.15), можно получить кинетическое уравнение

$$\frac{\partial f_n}{\partial t} = \frac{2}{\hbar} \text{Im} \sum_m \int \langle \delta f_{nm} \mathbf{E}(\mathbf{r}, t) \rangle \mathbf{D}_{nm}(\mathbf{r}) d^3 r \equiv I_n(t), \quad (3.44)$$

где введено обозначение $I_n(t)$ для интеграла столкновений. Подставляя (3.23) в (3.44), можно записать интеграл столкновений через спектральную плотность флуктуаций поля:

$$I_n = \frac{V}{3\hbar^2 (2\pi)^2} \sum_m \int d\omega d^3 k \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} \delta(\omega - \omega_{nm}) \times \\ \times \left\{ \left[(\delta \mathbf{E}^{\parallel} \delta \mathbf{E}^{\parallel})_{\omega \mathbf{k}} + (\delta \mathbf{E}^{\perp} \delta \mathbf{E}^{\perp})_{\omega \mathbf{k}} \right] (f_m - f_n) - \right. \\ \left. - 4\pi\hbar \text{Im} \left[-\frac{1}{\tilde{\epsilon}(\omega, \mathbf{k})} - \frac{2}{\tilde{\epsilon}(\omega, \mathbf{k}) - c^2 \mathbf{k}^2 / \omega^2} \right] (f_m + f_n) \right\}. \quad (3.45)$$

В отличие от выражений для энергии поля, интегрирование по частотам ω производится в пределах от $-\infty$ до ∞ . Первый член в квадратных скобках соответствует взаимодействию с продольным полем, в то время как второй — с поперечным полем. Поскольку соотношения (3.40), (3.41) удовлетворяются в равновесном состоянии, то, учитывая (3.42), получаем, что в равновесном состоянии интеграл столкновений I_n обращается в ноль.

Чтобы разделить вклады в интеграл столкновений, возникающие из-за индуцированных электрическим полем переходов от спонтанных переходов, необходимо вычесть вклад нулевых колебаний из объемной плотности энергии электромагнитного поля:

$$\pi \sigma_{\omega \mathbf{k}} = \frac{1}{4\pi} (\delta \mathbf{E} \delta \mathbf{E})_{\omega \mathbf{k}} - \\ - \hbar \text{Im} \left[-\frac{1}{\tilde{\epsilon}(\omega, \mathbf{k})} - \frac{2}{\tilde{\epsilon}(\omega, \mathbf{k}) - c^2 \mathbf{k}^2 / \omega^2} \right] \text{sign } \omega. \quad (3.46)$$

Подстановка плотности нулевых колебаний в (3.45) дает

$$\begin{aligned}
I_n &= \frac{V}{3\hbar^2\pi} \sum_m \int d\omega d^3k \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} \delta(\omega - \omega_{nm}) \times \\
&\times \left\{ \pi \sigma_{\omega\mathbf{k}} (f_m - f_n) - 2\hbar \operatorname{Im} \left[-\frac{1}{\tilde{\varepsilon}(\omega, \mathbf{k})} - \frac{2}{\tilde{\varepsilon}(\omega, \mathbf{k}) - c^2 \mathbf{k}^2 / \omega^2} \right] \times \right. \\
&\times \left. (f_m \theta(-\omega_{nm}) + f_n \theta(\omega_{nm})) \right\} \equiv \\
&\equiv \int d\omega d^3k \left\{ \sum_{m < n} \left[\sigma_{\omega\mathbf{k}} B_m^n (f_m - f_n) - A_m^n f_n \right] + \right. \\
&\left. + \sum_{m > n} \left[\sigma_{\omega\mathbf{k}} B_m^n (f_m - f_n) + A_m^n f_m \right] \right\}. \tag{3.47}
\end{aligned}$$

Здесь мы ввели коэффициенты Эйнштейна A_m^n и B_m^n , характеризующие соответственно спонтанные и индуцированные процессы:

$$\begin{aligned}
B_m^n &= \frac{1}{3\hbar^2} \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} \delta(\omega - \omega_{nm}), \\
A_m^n &= \frac{2}{3\hbar\pi} \left(|\mathbf{D}_{nm}|^2 \right)_{-\mathbf{k}} \delta(\omega - \omega_{nm}) \times \\
&\times \operatorname{Im} \left[-\frac{1}{\tilde{\varepsilon}(\omega, \mathbf{k})} - \frac{2}{\tilde{\varepsilon}(\omega, \mathbf{k}) - c^2 \mathbf{k}^2 / \omega^2} \right].
\end{aligned}$$

Порядок суммирования по n и m записан в предположении, что энергии E_n растут с ростом n . Индуцированные процессы вызываются как внешним электромагнитным полем, так и равновесными и неравновесными колебаниями поля.

Диэлектрическая проницаемость и плотность энергии электромагнитного поля в уравнениях (3.45) и (3.47) усредняются по ансамблю систем с функцией распределения f_n . Но эта функция распределения появляется в интегралах столкновения еще в одном месте явным образом. Независимое усреднение возможно только в случае, если точки излучения и поглощения виртуальных фотонов принадлежат различным статистически независимым подсистемам. Таким образом, в этом приближении вся система разбивается на подсистемы, которые могут взаимодействовать между собой только путем

обмена фотонами. Функция распределения f_n в этом случае является функцией распределения этих подсистем, а не для всей системы в целом.

Интеграл столкновений может быть также записан в форме интеграла столкновений Больцмана, если в (3.45) подставить выражения (3.32), (3.36) и (3.37):

$$I_n = \frac{4V\pi}{9\hbar^2} \sum_{mn'm'} \int d\omega d^3k \left(\mathbf{p}_{nm} \right)_{-\mathbf{k}}^2 \left(\mathbf{p}_{n'm'} \right)_{-\mathbf{k}}^2 \times \\ \times \delta(\omega - \omega_{nm}) \delta(\omega - \omega_{n'm'}) \left[\frac{1}{|\tilde{\epsilon}(\omega, \mathbf{k})|^2} + \frac{2}{|\tilde{\epsilon}(\omega, \mathbf{k}) - c^2 \mathbf{k}^2 / \omega^2|^2} \right] \times \\ \times (f_m f_{n'} - f_{m'} f_n) \}. \quad (3.48)$$

В такой формулировке поле исключается из интеграла столкновений, и процесс рассеяния (столкновения) более явным образом описывается в терминах обмена продольными (первый член в квадратных скобках) или поперечными (второй член) фотонами. Таким образом, система, например, переходит из состояния n в состояние m , испуская фотон с энергией $\hbar\omega = \hbar\omega_{nm}$ и импульсом $\hbar\mathbf{k}$, после чего этот фотон поглощается другой частью системы, переходящей из состояния m' в n' . Другие возможные процессы изображены на рис. 7.

Формулы (3.47) и (3.48) приводят к различным интерпретациям процессов, описываемых интегралом столкновений. Переходы, приводящие к испусканию и последующему поглощению фотонов, явным образом выделяются в форме (3.48) интеграла столкновений. Фотоны, участвующие в этих процессах, должны быть поглощены или испущены в рассматриваемой конденсированной среде, и не могут покинуть ее. С другой стороны, в интеграле столкновений в форме (3.47) происхождение поглощаемых фотонов (в членах, описывающих индуцированные процессы) или судьба испущенных фотонов (в членах, описывающих спонтанные и индуцированные процессы)

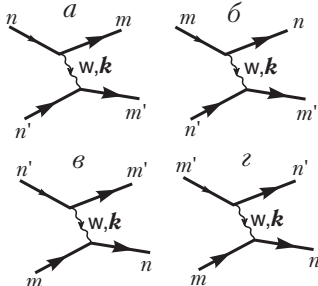


Рис. 7. Различные процессы, вносящие вклад в интеграл столкновений (3.48)

не конкретизируется. Тем самым интеграл столкновений в форме (3.47) удобно использовать при рассмотрении системы во внешнем электромагнитном поле или в случае, когда испущенные поперечные фотоны покидают рассматриваемую систему. Ниже мы будем использовать также и смешанную форму интеграла столкновений.

Если неравновесное состояние является анизотропным и неоднородным, необходимо использовать выражения для корреляторов компонент поля в форме (3.38) и выражения для нелокальной диэлектрической проницаемости (3.33). Это приводит к следующей форме интеграла столкновений:

$$I_n = \frac{2\pi}{\hbar^2} \sum_{mn'm'} \int d\omega \times \\ \times \left| \sum_{ik} \int d^3r d^3r_1 K_{ik}(\mathbf{r} - \mathbf{r}_1) (\mathbf{D}_{mn}(\mathbf{r}))_i (\mathbf{D}_{m'n'}(\mathbf{r}_1))_k \right|^2 \times (3.49) \\ \times \delta(\omega - \omega_{nm}) \delta(\omega - \omega_{n'm'}) (f_m f_{n'} - f_{m'} f_n) \}.$$

где $K_{ij}(\mathbf{r})$ определяется формулой (1.12). Такая форма удобна в том случае, если функции $|n\rangle$ локализованы в ограниченных пространственных областях, а переходы происходят также в локализованные состояния $|m\rangle$. В таком случае интеграл (3.49) будет описывать диполь-дипольный (для обмена продольными фотонами) и радиационный (в случае поперечных фотонов) перенос энергии между локализованными состояниями. Подробное рассмотрение этих вопросов будет проведено в других частях учебного пособия.

Глава 2

Состояния в недеформируемой решетке

§4 Электронные состояния в недеформируемой решетке

4.1 Иерархия приближений в теории твердого тела

В предыдущей главе было показано, что для получения конкретных результатов необходимо уметь вычислять волновые функции в твердом теле. Поскольку в одном кубическом сантиметре вещества находится порядка 10^{23} частиц, вычислить точно волновые функции невозможно. При изучении твердого тела широко используются различные приближения, что соответствует различным представлениям гамильтониана H_0 . В связи с этим сделаем несколько общих замечаний об иерархии приближений, используемых в физике твердого тела.

Для описания различных явлений, связанных с взаимодействием света с веществом, необходимо использовать разные приближения. Так, при описании поглощения света при больших энергиях фотонов, в глубине области фундаментального поглощения, часто оказывается достаточно применить простейшее приближение — описывать невзаимодействующие электроны в поле недеформируемой решетки из ионных остовов. Однако для описания явлений на краю фундаментального поглощения необходимо учитывать взаимодействие между электронами. Это дает качественно правильное

описание спектра поглощения. Для того чтобы описать форму линии краевого поглощения, приближение о недеформируемой кристаллической решетке уже не пригодно, и необходимо пользоваться представлением о фононах — квантах колебаний решетки. В области примесного поглощения приходится отказываться от идеальности кристалла. Форма же линии поглощения примеси в принципе не может быть описана в приближении недеформируемой решетки. В области сильного поглощения в ряде твердых тел нельзя даже пользоваться рассмотрением взаимодействия с фотонами в рамках теории возмущения. Это приводит к представлениям о поляритоне.

Подобный список различных приближений может быть продолжен. Естественно, что для описания каждого явления необходимо ограничиться простейшим приближением, адекватно описывающим характерные черты данного явления. Такое разнообразие приближений в твердом теле приводит к ряду методических трудностей, когда более сложное приближение строится практически в отрыве от предыдущего, а их взаимосвязь часто опускается. Кроме того, в различных приближениях используются несколько различным языком. Это ведет к тому, что близкие явления называются различными терминами, например возбужденное состояние примеси и экситон, захваченный примесью, с соответствующей разницей в интерпретации физических результатов.

Расчету электронных свойств твердых тел посвящено большое количество прекрасных книг (см., например, [7, 8, 56–58] и др.), и здесь не будут обсуждаться все связанные с такими расчетами проблемы. Поскольку основное внимание в данной книге уделяется спектроскопии твердых тел, мы ограничимся введением только основных понятий.

4.2 Одноэлектронное приближение

В соответствии с вышесказанным начнем описание твердого тела с простейшего приближения. Гамильтониан твердого тела в нерелятивистском пределе имеет вид

$$\hat{H}_0 = \sum_I \frac{\hat{\mathbf{P}}_I^2}{2M_I} + \sum_i \frac{\hat{\mathbf{p}}_i^2}{2m_e} + \sum_{I < J} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{i, I} \frac{Z_I e^2}{|\mathbf{R}_I - \mathbf{r}_i|} + \sum_{i < j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \hat{H}_s . \quad (4.1)$$

Здесь первый член описывает кинетическую энергию ядер, второй — электронов. Третий, четвертый и пятый члены описывают кулоновское взаимодействие соответственно между ядрами и ядрами, между электронами и ядрами и между электронами и электронами. В последний член, символически записанный как $\hat{\mathcal{H}}_s$, включены все спиновые эффекты. В основном в этот член вносит вклад спин-орбитальное взаимодействие для электронов.

Этот гамильтониан описывает практически все эффекты в твердом теле. Однако в такой записи пользоваться им бессмысленно — найти собственные функции $\hat{\mathcal{H}}_0$ не представляется возможным. Поэтому в качестве первого шага попытаемся максимально упростить выражение (4.1). Воспользуемся тем, что масса ядер M более чем на три порядка отличается от массы электронов m . В нулевом приближении можно считать ядра неподвижными (этим приближением мы будем пользоваться до конца данной главы). Это позволяет опустить первый и третий члены: первый член мал по сравнению с кинетической энергией электронов, а третий приводит лишь к постоянному сдвигу начала отсчета энергии. В таком случае $\hat{\mathcal{H}}_0$ зависит лишь от электронных координат и импульсов:

$$\hat{\mathcal{H}}_0 = \sum_i \frac{\hat{\mathbf{p}}_i^2}{2m_e} + \sum_i U(\mathbf{r}_i) + \sum_{i < j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \hat{\mathcal{H}}_s . \quad (4.2)$$

В идеальном кристалле потенциальная энергия электрона в поле ядер $U(\mathbf{r}_i)$ является периодической функцией \mathbf{r} с периодами, определяемыми решеткой Браве кристалла. Однако и после такого упрощения решить задачу на собственные функции не представляется возможным. Практически всегда эту задачу решают в так называемом «одноэлектронном» приближении, когда волновая функция аппроксимируется детерминантом Слетера, построенным из функций, каждая из которых зависит от координат только одного электрона. Число таких функций и, следовательно, порядок детерминанта совпадают с числом электронов в твердом теле. Соответствующее приближение в теории многоэлектронных систем называется приближением Хартри — Фока. При этом многоэлектронная задача сводится к одноэлектронной, но с частичным учетом электрон-электронного взаимодействия.

Собственные функции гамильтониана идеального

твердого тела в одноэлектронном приближении определяются квантовыми числами — волновым вектором \mathbf{k} и номером зоны s . В периодическом кристалле понятие волнового вектора довольно сильно отличается от понятия волнового вектора электрона в свободном пространстве. Для того, чтобы познакомиться с основными свойствами волновых векторов в кристалле, необходимо оказать несколько слов о теории групп. Об этой абстрактной математической теории написано много превосходных монографий и учебников. Применительно к проблемам физики твердого тела аппарат теории групп изложен, например, в [4, 49] и ряде других учебников, целиком посвященных теории групп (например, [59]). Теория групп, исходя из свойств симметрии физической системы, позволяет систематизировать кристаллы по их симметрии, классифицировать электронные состояния как идеального кристалла, так и кристалла с примесным атомом, определить правила отбора для матричных элементов перехода в системе под действием внешних возмущений. За последовательным изложением теории групп мы отсылаем читателя к вышеупомянутым учебникам, а здесь на практически важных примерах рассмотрим некоторые следствия, широко применяемые в спектроскопии.

Интересующие нас группы симметрии кристалла являются совокупностями операций симметрии, совмещающих кристалл сам с собой. При этом кристалл рассматривается либо как бесконечный, либо на него налагаются циклические граничные условия Борна-Кармана, согласно которым все волновые функции считаются периодическими с периодами, много большими периодов решетки кристалла и кратными последним. В одномерном случае циклические условия Борна-Кармана эквивалентны замыканию длинной цепочки атомов в кольцо. Совокупность всех трансляций кристалла на векторы

$$\mathbf{n} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3, \quad (4.3)$$

где $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ — основные векторы трансляций кристалла, а n_i — целые числа, образуют группу трансляций кристалла. Кроме трансляций существуют операции симметрии, переводящие эквивалентные векторы в кристалле друг в друга. Эти операции симметрии образуют точечную группу кристаллической решетки. В точечную группу симметрии могут входить: инверсия, повороты вокруг оси на 60° и 90° и кратные им углы; отражения в плоскостях и некоторые комбинации из этих элементов. Точечные группы кристаллов всегда

содержат конечное число элементов. Так, одна из самых «богатых» групп точечной симметрии — группа O_h (группа симметрии куба) — содержит 48 элементов. Полная группа кристалла, так называемая пространственная группа, содержит набор совокупностей элементов трансляций и элементов точечной симметрии. Некоторые группы (называемые несимморфными в отличие от остальных, именуемых симморфными) содержат так называемые неэлементарные трансляции, то есть сдвиг на половину периода совместно с отражением в плоскости (плоскости скольжения), или сдвиг на одну треть периода совместно с поворотом вокруг некоторой оси на 120° (винтовые оси).

Теория групп позволила провести полную классификацию кристаллов. Все кристаллы относятся к семи системам или сингониям: 1) триклинной, 2) моноклинной, 3) ортогональной, или ромбической, 4) квадратичной, или тетрагональной, 5) гексагональной, 6) тригональной, или ромбоэдрической, 7) кубической. В зависимости от типа решетки — простой, базоцентрированной, объемноцентрированной или гранецентрированной — эти семь систем подразделяются на 14 типов решеток. Внутри этих типов можно выделить 230 пространственных групп. Тем самым, например, большинство щелочно-галлоидных кристаллов, обладающих одной и той же пространственной группой, с точки зрения симметрии совершенно идентичны.

Все применения теории групп в квантовой механике основаны на теории представлений групп. Из того факта, что кристаллическая решетка кристалла под действием операций симметрии совмещается сама с собой, не следует, что волновые функции ведут себя так же. В простейшем случае они под действием операторов симметрии домножаются на некоторый фазовый множитель, в более сложных случаях выражаются через линейную комбинацию исходных функций. В системе волновых функций можно выделить базис функций. Под действием операций симметрии \hat{P}_S каждая базисная функция ψ_i преобразуется в линейную комбинацию базисных функций:

$$\hat{P}_S \psi_i = \sum_k \Gamma(S)_{ki} \psi_k.$$

Коэффициенты преобразования Γ_{ki} образуют матрицу для каждого элемента группы S . Совокупность этих матриц для

всех элементов группы называется представлением группы.

Представление $\Gamma(S)_{ki}$ называется приводимым, если с помощью унитарных преобразований можно так преобразовать базис $\{\psi_i\}$, что все элементы представления $\Gamma(S)_{ki}$ для всех элементов S группы принимают вид

$$\Gamma(S) = \begin{pmatrix} \Gamma'(S) & 0 \\ 0 & \Gamma''(S) \end{pmatrix},$$

причем волновые функции каждой из частей Γ' и Γ'' преобразуются через функции своей же части. Сумма размерностей матриц Γ' и Γ'' равна размерности матрицы Γ . В противном случае представление называется неприводимым. Неприводимые представления играют основную роль в применениях теории групп в физике. Это связано с тем, что волновые функции, соответствующие определенному уровню энергии, преобразуются по неприводимому представлению.

В качестве примера неприводимых представлений можно привести атомные функции s -, p -, d - и т. д. При этом s -функции образуют одномерное неприводимое представление группы вращений вокруг центра атома на произвольные углы, p -функции — трехмерное неприводимое представление и т. д. При наложении магнитного поля симметрия системы понижается — остаются только вращения вокруг направления магнитного поля. В этой новой группе все многомерные p -, d - и т. д. представления оказываются уже приводимыми. Физически это выливается в расщеплении $(2l + 1)$ -кратно вырожденно уровня на $2l + 1$ подуровней.

В качестве другого примера рассмотрим волновую функцию свободного электрона $V^{-1/2} \exp(i\mathbf{k}\mathbf{r})$ (V является нормировочным объемом системы), которая преобразуется по неприводимому представлению группы трансляций на произвольные векторы. Под действием трансляции на вектор \mathbf{a} волновая функция умножается на $\exp(i\mathbf{k}\mathbf{a})$. Если в первом примере номером неприводимого представления являлось орбитальное квантовое число l , то в последнем случае — волновой вектор \mathbf{k} , который может принимать произвольные значения.

Вернемся к пространственной группе кристалла. Неприводимые представления группы трансляций кристалла тоже классифицируются волновым вектором \mathbf{k} . Однако теперь волновой вектор может принимать значения только в ограниченной области — так называемой первой зоне Бриллюэна (

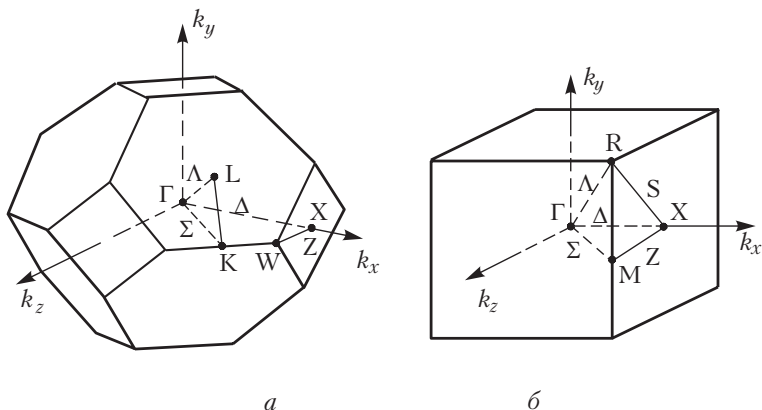


Рис. 8. Зоны Бриллюэна для ГЦК (а) и простой кубической решетки (б). Приведены обозначения высокосимметричных точек.

слово «первая» часто опускается). В одномерном случае первой зоной Бриллюэна является отрезок $(-\pi/a, \pi/a)$, где a — период одномерной цепочки. Форма зоны Бриллюэна в трехмерном случае часто является довольно сложной. На рис. 8 приведены зоны Бриллюэна для кристаллов с гранецентрированной и простой кубической решеткой. Для удобства обозначений ряд высокосимметричных точек зоны Бриллюэна именуется греческими или латинскими буквами. Центр зоны Бриллюэна (точка $\mathbf{k} = 0$) всегда называется Γ -точкой. Обозначения ряда других точек симметрии приведены на рис. 8. Обычно внутренние точки, соответствующие осям высокой симметрии, обозначаются греческими буквами (Λ , Δ , Σ и т. д.), а точки, лежащие на границах зоны Бриллюэна — латинскими (X , K , L и т. д.).

Можно показать, что в одноэлектронном приближении волновая функция электрона в кристалле имеет вид

$$\psi_{\mathbf{k}s}(\mathbf{r}) = V^{-1/2} u_{\mathbf{k}s}(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}}. \quad (4.4)$$

Здесь $u_{\mathbf{k}s}(\mathbf{r}) = u_{\mathbf{k}s}(\mathbf{r} + \mathbf{n})$ — пространственно-периодическая функция с периодами, равными векторам основных трансляций (4.3). Функция (4.4) называется функцией Блоха. Конкретный вид функции $u_{\mathbf{k}s}(\mathbf{r})$ определяется гамильтонианом

$\hat{\mathcal{H}}_0$. Блоховская волновая функция характеризуется не только волновым вектором \mathbf{k} , но и номером зоны s (аналогичным главному квантовому числу в атоме).

В произвольных точках зоны Бриллюэна уровни энергии не вырождены, т. е. на каждом уровне может находиться по два электрона с противоположно направленными спинами. Однако в точках высокой симметрии возможно вырождение, определяемое пространственной группой кристалла. Для большинства групп максимальная кратность вырождения равна трем. При этом соответствующие электронные состояния чаще всего нумеруют согласно нумерации неприводимых представлений. Так, в Γ -точке в гранецентрированных кубических (ГЦК) кристаллах (например, в NaCl) возможны однократно вырожденные состояния Γ_1 и Γ_2 , двухкратно вырожденные состояния Γ_{12} и трехкратно вырожденные Γ_{15} и Γ_{25} . В зависимости от поведения функций под действием операции инверсии обозначения соответствующих состояний могут иметь дополнительный значок « $-$ » для нечетных по отношению к инверсии волновых функций. К сожалению, общая система обозначений довольно сложна. С ней можно познакомиться в работе [60]. При уходе из точки $\mathbf{k} = 0$ большинство состояний расщепляется. Однако, например, если волновой вектор \mathbf{k} остается на оси высокой симметрии, например Δ , снятое вырождения может быть неполным. Так, на оси Δ возможны однократно вырожденные состояния Δ_1 , Δ_2 и двухкратно вырожденное Δ_5 . Можно вывести так называемые соотношения совместности, которые показывают, что, например, Γ_1 и Γ_2 -состояния переходят в Δ_1 и Δ_2 -состояния соответственно, Γ_{12} расщепляется на Δ_1 и Δ_2 -состояния, Γ_{15} — на Δ_1 и Δ_5 , Γ_{25} — на Δ_2 и Δ_5 (отсюда, кстати, получается определенное обособление индексации Γ -состояний).

Поскольку изобразить зависимость энергии состояний от волнового вектора во всей зоне Бриллюэна не представляется возможным, обычно ограничиваются изображением функциональной зависимости $E(\mathbf{k})$ только вдоль определенных направлений в зоне Бриллюэна.

Приведенная выше классификация состояний справедлива для случая, когда не учитывается спин-орбитальное взаимодействие. При учете этого взаимодействия происходит дополнительное расщепление электронных состояний. Так, шестикратно (при учете двух возможных положений спина

электрона) вырожденное состояние Γ_{25} распадается на двукратно вырожденное Γ_7 и четырехкратно вырожденное Γ_8 -состояния, соответствующие неприводимым представлениям так называемых двойных групп (учитывающих наличие спиновой переменной). Индексы в этом случае уже не имеют наглядной интерпретации.

4.3 Приближение почти свободных электронов

Прежде чем рассматривать общий случай, посмотрим, к чему приводит периодичность решетки в одномерном случае в двух пределах. Во-первых, рассмотрим так называемую модель почти свободных электронов. Закон дисперсии для свободных электронов имеет вид $E(k) = \hbar^2 k^2 / 2m_e$, причем k в этой области изменяется в бесконечных пределах. Волновая функция имеет вид $V^{-1/2} e^{ikx}$. Теперь введем слабый периодический потенциал $U(x) = U(x + a)$ и попытаемся привести бесконечную область изменения k к первой зоне Бриллюэна $(-\pi/a, \pi/a)$. Если $|k| < \pi/a$, то волновой вектор попадает в первую зону Бриллюэна, и в нулевом приближении волновая функция имеет вид функции Блоха с $u_{k,0}(x) = 1$. Если волновой вектор свободного электрона находится в пределах второй зоны Бриллюэна (например, $\pi/a < k' < 2\pi/a$), то необходимо сдвинуть дисперсионную кривую, введя квазиволновой вектор $k = k' - 2\pi/a$. При таком сдвиге квазиволновой вектор попадает в первую зону Бриллюэна, а соответствующая ветвь дисперсионной кривой имеет индекс «1». Волновая функция свободного электрона может быть приведена к функции Блоха после выделения периодической функции $u_{k,1}(x) = e^{2i\pi x/a}$:

$$V^{-1/2} e^{ik'x} = V^{-1/2} u_{k,1}(x) e^{ikx}.$$

Эта операция называется приведением к первой зоне Бриллюэна и может быть продолжена для всех волновых векторов. В результате получаем зависимость энергии от волнового вектора, показанную пунктирной кривой на рис. 9. Функция $E(k)$ становится многозначной функцией, и ее ветви нумеруются от нуля до бесконечности.

Такая картина справедлива только в предположении пренебрежения потенциалом $U(x)$. В первом приближении по периодическому возмущению возникнут расщепления пересекающихся уровней, и зависимость энергии от волнового вектора

окончательно приобретет вид, показанный на рис. 9 сплошной линией. Слабый потенциал изменяет только области k вблизи пересекающихся уровней. Расщепление пропорционально квадратному корню из матричного элемента потенциала $U(x)$ и даже для слабых потенциалов не является пренебрежимым.

Это расщепление состояний создает запрещенные зоны в спектре энергий, то есть зоны, в которых нет решений, соответствующих бесконечной системе. Необходимо отметить, что в приближении почти свободных электронов такие запрещенные зоны возникают только в одномерном случае. В дву- и трехмерных случаях аналогичное расщепление возникает на границах зоны Бриллюэна и в некоторых точках высокой симметрии в зоне Бриллюэна, но при этом щели в различных высокосимметричных точках сдвинуты друг относительно друга, и поэтому зоны разрешенных энергий в одних направлениях перекрываются с зонами запрещенных энергий в других направлениях в зоне Бриллюэна. Дисперсия состояний для ГЦК-решетки с параметрами, соответствующими твердому ксенону, приведена на рис. 10.

Закон дисперсии $E(k) = \hbar^2 k^2 / 2m_e$ может быть переписан в виде, более удобном для проведения оценок: $E(k) = Ry (ka_B)^2$, где введена постоянная Ридберга для атома водорода $Ry = \hbar^2 / 2m_e a_B^2 = 13,6 \text{ эВ}$, а $a_B = \hbar^2 / m_e e^2 = 0,053 \text{ нм}$ — боровский радиус. Пользуясь этой формулой легко оценить энергию, при которой возникает первое пересечение волнового вектора с границей зоны Бриллюэна $k = \pi/a$: $E(\pi/a) = Ry (\pi a_B / a)^2$ и для постоянной решетки $a = 0,25 \text{ нм}$ имеет значение порядка 6 эВ.

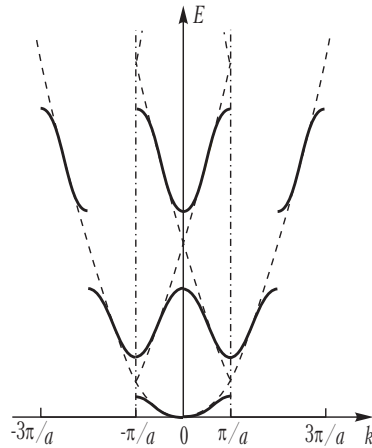


Рис. 9. Зависимость энергии от волнового вектора для одномерного случая в модели свободных электронов (пунктир) и почти свободных электронов (сплошные линии).

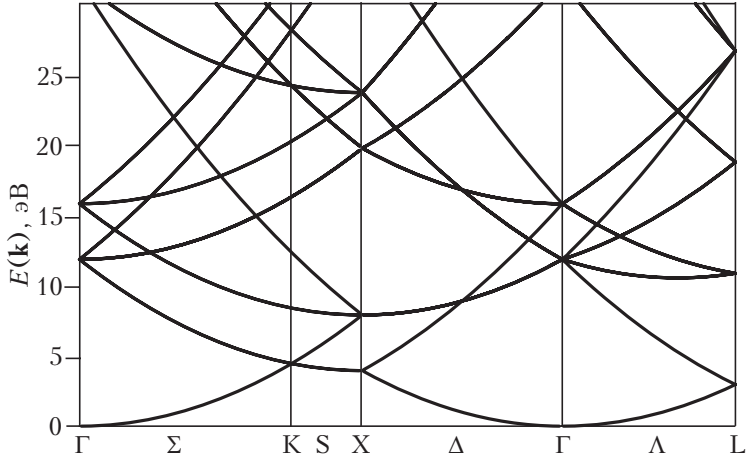


Рис. 10. Закон дисперсии электронов для почти свободных электронов в ГЦК-решетке с постоянной решеткой, соответствующей твердому ксенону.

Приближение слабо связанных электронов имеет не только методическое значение, но оно адекватно описывает высоколежащие состояния в зоне проводимости, энергия которых больше 10–30 эВ. Более того, для таких электронных состояний периодическим потенциалом можно часто пренебречь, и такие электроны могут рассматриваться как свободные (с массой m_e , равной массе свободных электронов).

4.4 Приближение сильно связанных электронов

Другой простейший пример — приближение сильно связанных электронов. Оно пригодно для описания остовных электронов. В простейшем случае предполагается, что окружающие атомы и ионы слабо воздействуют на электронное состояние выделенного атома. Одноэлектронная волновая функция φ_α изолированного атома, соответствующая состоянию α с энергией ε_α , удовлетворяет уравнению

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + w(\mathbf{r}) - \varepsilon_\alpha \right] \varphi_\alpha(\mathbf{r}) = 0,$$

в которое входит потенциал иона $w(\mathbf{r})$, действующий на

рассматриваемый электрон (при этом пренебрегаем всеми корреляционными и обменными эффектами).

В первом приближении можно пренебречь перекрыванием волновых функций, относящихся к разным узлам решетки, находящимся в точках \mathbf{n} , \mathbf{n}' :

$$\int \varphi_{\alpha}^*(\mathbf{r} - \mathbf{n}) \varphi_{\beta}(\mathbf{r} - \mathbf{n}') d^3 r = \delta_{\alpha\beta} \delta_{\mathbf{nn}'},$$

где \mathbf{n} и \mathbf{n}' определены в (4.3). Функцию первого приближения, удовлетворяющую теореме Блоха, можно искать в виде

$$\Psi_{\mathbf{k}\alpha} = N^{-1/2} \sum_{\mathbf{n}} e^{i\mathbf{k}\mathbf{n}} \varphi_{\alpha}(\mathbf{r} - \mathbf{n}),$$

где N — число атомов в кристалле (для простоты кристалл предполагается одноатомным). Гамильтониан электрона в кристалле отличается от гамильтониана одного атома наличием потенциальных энергий от соседних атомов:

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \nabla^2 + \sum_{\mathbf{n}} \omega(\mathbf{r} + \mathbf{n}).$$

При этом энергия электрона в первом приближении имеет вид

$$E_{\alpha}(\mathbf{k}) = \varepsilon_{\alpha} + \sum_{\mathbf{n}} e^{i\mathbf{k}\mathbf{n}} \int \varphi_{\alpha}^*(\mathbf{r} + \mathbf{n}) \sum_{\mathbf{m} \neq 0} \omega(\mathbf{r} - \mathbf{m}) \varphi_{\alpha}(\mathbf{r}) d^3 r.$$

Часто можно ограничиться приближением ближайших соседей, то есть ограничить суммирование по \mathbf{n} только первой координационной сферой (в то же время в сумме по \mathbf{m} можно оставить все члены, поскольку потенциал $\omega(\mathbf{r})$, имеющий на больших расстояниях дальнедействующий характер, спадает существенно медленнее, чем атомные волновые функции $\varphi_{\alpha}(\mathbf{r})$). В этом случае для простой кубической решетки закон дисперсии имеет следующий вид:

$$E_{\alpha}(\mathbf{k}) = \varepsilon_{\alpha} - w_0 - 2w_a [\cos k_x a + \cos k_y a + \cos k_z a], \quad (4.5)$$

где

$$w_0 = -\int \varphi_{\alpha}^*(\mathbf{r}) \sum_{\mathbf{m} \neq 0} \omega(\mathbf{r} - \mathbf{m}) \varphi_{\alpha}(\mathbf{r}) d^3 r > 0,$$

$$w_a = -\int \varphi_{\alpha}^*(\mathbf{r} + \mathbf{a}) \sum_{\mathbf{m} \neq 0} \omega(\mathbf{r} - \mathbf{m}) \varphi_{\alpha}(\mathbf{r}) d^3 r.$$

Величина w_0 определяет сдвиг середины зоны по отношению к основному уровню изолированного атома, а $12|w_a|$ — ширина зоны разрешенных энергий.

Для длинноволновых состояний ($ka \ll 1$) можно разложить $E_\alpha(\mathbf{k})$ в ряд по ka :

$$E_\alpha(\mathbf{k}) = \varepsilon_\alpha - w_0 - 6w_a + \frac{\hbar^2 k^2}{2m^*},$$

где введено понятие эффективной массы $m^* \equiv \hbar^2/2a^2w_a$. Знак эффективной массы при $\mathbf{k} = 0$ зависит от знака w_a , т. е. она может быть и отрицательной. Эффективная масса электронного возбуждения имеет в квазиклассическом пределе смысл массы во втором законе Ньютона для электронного возбуждения. Это понятие применимо только в непосредственной близости от экстремума ($ka \ll 1$). В случае отрицательной эффективной массы отрицательно заряженный электрон ведет себя в электромагнитном поле как положительно заряженная частица. Положительная эффективная масса соответствует дну зоны разрешенных энергий, а отрицательная — потолку зоны. Поскольку приближение сильно связанных электронов справедливо только при малом перекрытии волновых функций, величина w_a должна быть относительно малой. Вследствие этого данное приближение справедливо только для узких зон, и поэтому эффективные массы электрона в таких зонах много больше массы свободного электрона.

Выражение (4.5) достигает экстремума и во всех углах зоны Бриллюэна, то есть в точках $|k_x| = |k_y| = |k_z| = \pi/a$. Вблизи этих экстремумов закон дисперсии имеет вид

$$E_\alpha(\mathbf{k}) = \varepsilon_\alpha - w_0 + 6w_a + \hbar^2 \frac{(|k_x| - \pi/a)^2 + (|k_y| - \pi/a)^2 + (|k_z| - \pi/a)^2}{2m^{**}},$$

где $m^{**} = -m^* = -\hbar^2/2a^2w_a$.

Кроме этих двух экстремумов, зависимость энергии от импульса (4.5) имеет несколько стационарных точек, в которых обращаются в ноль производные $\partial E_\alpha / \partial \mathbf{k}$. Например, в точке $k_x = 0, k_y = k_z = \pi/a$ имеем

$$E_\alpha(\mathbf{k}) = \varepsilon_\alpha - w_0 + 6w_a + \frac{\hbar^2 k_x^2}{2m^*} + \frac{\hbar^2 (k_y - \pi/a)^2}{2m^{**}} + \frac{\hbar^2 (k_z - \pi/a)^2}{2m^{**}}.$$

Поскольку m^* и m^{**} имеют противоположные знаки, энергетические поверхности вблизи от таких точек представляют из себя гиперболоиды. Такие стационарные точки, для которых эффективные массы по разным направлениям имеют различные знаки, называются седловыми точками. В трехмерном случае седловые точки могут быть двух типов: с одной отрицательной и двумя положительными массами, и с двумя положительными и одной отрицательной массой. В окрестности седловой точки движение электрона достаточно необычно: если электрическое поле приложено в одном направлении, электрон ведет себя как отрицательно заряженная частица, а в другом направлении – как положительно заряженная.

На рис. 11 изображены изоэнергетические поверхности в двумерной зоне Бриллюэна для прямоугольной решетки. Для энергий вблизи дна зоны (мы рассматриваем случай $m^* > 0$), то есть для малых \mathbf{k} , изоэнергетические поверхности являются замкнутыми эллипсами вокруг Γ -точки (в трехмерном случае они являются эллипсоидами или, для кубической решетки – сферами). С ростом энергии изоэнергетические поверхности деформируются, и при энергии, равной энергии одной из седловых точек, эти поверхности начинают пересекать границу зоны Бриллюэна. Изоэнергетические поверхности из замкнутых превращаются в открытые, то есть меняется их топология. В схеме расширенных зон Бриллюэна они представляют из себя разнообразные периодические структуры: гофрированные цилиндры и плоскости, а также другие структуры достаточно общего вида. С дальнейшим ростом энергии топология этих изоэнергетических поверхностей меняется по крайней мере еще раз. При последней смене топологии поверхности становятся опять замкнутыми. На рис. 11 четыре части

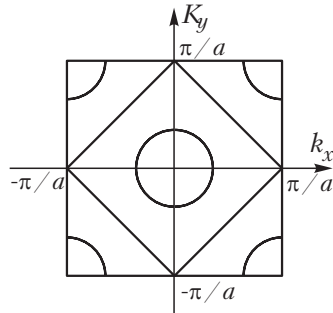


Рис. 11. Изоэнергетические поверхности в простой двумерной кубической решетке в приближении сильной связи.

эллипсов около угловых точек зоны Бриллюэна как раз и соответствуют таким замкнутым поверхностям. Из-за периодичности в пространстве квазиимпульса эти части эллипса могут быть объединены в один эллипс вокруг любого из углов зоны Бриллюэна.

Упомянутые выше изменения топологии изоэнергетических поверхностей с ростом энергии от дна зоны до потолка вовсе не являются следствием модели сильной связи, но являются типичными для всех типов одноэлектронного приближения. Это утверждение о числе точек изменения топологии составляет содержание теоремы Ван Хофа [61]. Изменение топологии происходит также и на потолке и дне зоны, где изоэнергетические поверхности исчезают или появляются, соответственно, с ростом энергии. Тем самым, с каждой зоной связаны как минимум четыре изменения топологии изоэнергетических поверхностей.

Выше рассматривался лишь случай, когда состояние атома, формирующее зону, не является вырожденным. При наличии вырождения (например, если зона формируется из p или d -состояний атома) уравнение (4.5) превращается в систему уравнений, и для определения энергии можно записать секулярное уравнение. Например, в случае формирования зоны из p -состояний для кристалла ГЦК-структуры можно получить [62]:

$$\begin{vmatrix} H_1(x, y, z) - E & H_2(y, x) & H_2(z, x) \\ H_2(x, y) & H_1(y, z, x) - E & H_2(z, y) \\ H_2(x, z) & H_2(y, z) & H_1(z, x, y) - E \end{vmatrix} = 0,$$

$$H_1(x, y, z) = (\omega_p - \omega_\pi)[\cos(x + y) + \cos(x - y) + \cos(x + z) + \cos(x - z)] - 2\omega_\pi[\cos(y + z) + \cos(y - z)],$$

$$H_2(x, y) = (\omega_p + \omega_\pi)[\cos(x + y) - \cos(x - y)],$$

$$x = k_x d, \quad y = k_y d, \quad z = k_z d.$$

Здесь ω_p и ω_π – интегралы типа для p -функций, ориентированных вдоль оси, соединяющей атомы, и перпендикулярно ей, соответственно, а $2d$ – постоянная решетки. Это кубическое уравнение имеет простые решения для высокосимметричных направлений. Зависимость $E(\mathbf{k})$ показана на рис. 12. Такого типа зависимость характерна для верхней заполненной зоны в щелочногалоидных соединениях. Видно, что

состояния на потолке зоны в Γ -точке являются трехкратно вырожденными и образуют подзоны легких и тяжелых дырок.

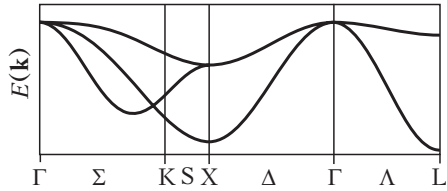
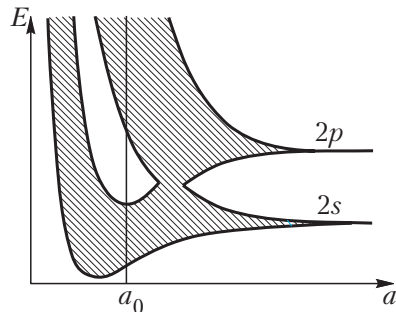


Рис. 12. Дисперсионные кривые в приближении сильной связи для зоны, сформированной p -функциями атомов (ГЦК-решетка)

На примере приближения сильной связи можно показать образование энергетических зон при сближении системы атомов. Дело в том, что ширина энергетической зоны $12|w_a|$ зависит от межатомного расстояния, экспоненциально увеличиваясь с уменьшением последнего. При этом можно нарисовать зависимость разрешенных зон энергий как функцию постоянной решетки a гипотетической системы атомов (рис. 13). Предположим, что атомная система имеет один заполненный и один свободный уровень. При сближении атомов за счет изменения w_0 и w_a с расстоянием каждый из уровней расширяется при одновременном изменении их положения. Приближение сильной связи, во всяком случае, перестает работать, когда получившиеся зоны энергий начинают пересекаться. После пересечения уже нельзя сказать, что нижняя зона сформирована из нижнего состояния атома, а верхняя — из верхнего. Каждая из зон построена на суперпозиции атомных функций обоих состояний. Такое явление как раз характерно для верхних валентных зон и нижних зон проводимости многих кристаллов. Это приводит к необходимости сильно увеличивать базис атомных функций для расчета энергетических спектров.



В общем случае задача расчета электронных состояний в твердом теле — чрезвычайно громоздкая задача, которая может быть решена только на больших ЭВМ, да и то в достаточно грубых приближениях. Большая часть из этих приближений

Рис. 13. Изменение разрешенных зон энергий в зависимости от расстояния между атомами в алмазе.

в своей основе содержит две рассмотренные выше модели. Так, в приближении присоединенных плоских волн (ППВ, APW) и ряде других весь объем кристалла разбивается на две части: на сферы вокруг атомов и пространство между атомными сферами (такой способ называется *muffin-tin*, по названию формы для выпечки сдобы). В первой области в качестве базисных используются атомные или ионные функции, во второй — волновые функции свободных электронов. Эти функции сшиваются на границе сфер. Поэтому приближенная волновая функция электрона внутри кристалла оказывается связанной с определенными атомными состояниями (или состояниями ионов в ионных кристаллах). В настоящее время в связи со все большим развитием вычислительной техники появилось достаточно много современных подходов, которые основаны как на так называемых первопринципных (*ab initio*) методах расчета, так и с использованием псевдопотенциалов, с различной степенью учета многочастичных эффектов. С помощью этих подходов можно рассчитывать зонные структуры и волновые функции в твердых телах с достаточно высокой степенью точности.

В сложных кристаллах, содержащих большое число атомов в элементарной ячейке, часто основную роль играют волновые функции группы атомов (так называемого кластера), которые могут быть исследованы независимо от остального объема кристалла. Поэтому до сих пор при изучении сложных систем, которые не могут быть рассчитаны как кристаллы, применяются расчеты молекулярных состояний отдельных групп атомов.

4.5 Плотность зонных состояний в твердом теле

Для характеристики энергетического спектра твердого тела, кроме зависимости энергии от волнового вектора $E_s(\mathbf{k})$, можно использовать плотность состояний $g(E)$, которая является числом различных состояний, попадающих в единичный интервал энергии около E :

$$g(E) = \sum_s g_s(E), \quad g_s(E) = \frac{V}{(2\pi)^3} \int_{\Omega_B} \delta(E_s(\mathbf{k}) - E) d^3k, \quad (4.6)$$

где Ω_B — объем зоны Бриллюэна, V — объем системы. Интеграл в данной формуле можно преобразовать в интеграл по элементам $dS_{\mathbf{k}}$ изоэнергетической поверхности в

в \mathbf{k} -пространстве и по энергии. Выполняя затем интегрирование по энергии (с учетом δ -функции в (4.6)), получаем

$$g_s(E) = \frac{V}{8\pi^3} \oint_{E_s(\mathbf{k})=E} \frac{dS_{\mathbf{k}}}{|\nabla_{\mathbf{k}} E_s(\mathbf{k})|}. \quad (4.7)$$

Около экстремума или стационарной точки, которая достигается при $\mathbf{k} = \mathbf{k}_{\text{ext}}$, энергию $E_s(\mathbf{k})$ можно разложить в ряд. В общем случае кристалла произвольной симметрии для невырожденной зоны получаем

$$E_s(\mathbf{k}) = E_s(\mathbf{k}_{\text{ext}}) + \sum_{i,j=x,y,z} \left(\frac{1}{m} \right)_{ij} (\mathbf{k} - \mathbf{k}_{\text{ext}})_i (\mathbf{k} - \mathbf{k}_{\text{ext}})_j. \quad (4.8)$$

После приведения к главным осям тензор обратной эффективной массы $(m^{-1})_{ij}$ может иметь различное число (от 0 до 3) отрицательных компонент. В первом случае (который обозначается как M_0), энергия имеет минимум, а в последнем случае (M_3) - максимум. Промежуточные случаи (M_1 и M_2) соответствуют особенностям типа седловых точек. В соответствии с этим плотность состояний имеет не менее четырех особенностей для каждой зоны, связанные с изменением топологии изоэнергетических поверхностей. Каждая из этих особенностей приводит к неаналитическим добавкам к плотности состояний, которые называются особенностями Ван Хова. Эти особенности впервые обсуждались Ван Ховом [61] для колебаний решетки. В трехмерном случае эти добавки имеют вид

$$\begin{aligned} M_0: \quad \Delta g_s(E) &\sim \sqrt{E - E_{\text{ext}}} \theta(E - E_{\text{ext}}), \\ M_1: \quad \Delta g_s(E) &\sim -\sqrt{E_{\text{ext}} - E} \theta(E_{\text{ext}} - E), \\ M_2: \quad \Delta g_s(E) &\sim -\sqrt{E - E_{\text{ext}}} \theta(E - E_{\text{ext}}), \\ M_3: \quad \Delta g_s(E) &\sim \sqrt{E_{\text{ext}} - E} \theta(E_{\text{ext}} - E), \end{aligned} \quad (4.9)$$

$$\text{где } \theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Тем самым вклад от каждой зоны в плотность состояний имеет вид, схематически показанный на рис. 14.

Плотность состояний для зоны, построенной из вырожденных функций, имеет большое число особенностей Ван Хова.

k

В двумерном и одномерном случаях особенности Ван Хова выражены ярче. Например, в двумерном случае $\sqrt{\Delta E}$ в формулах (4.9) заменяется либо на $\ln \Delta E$, либо изменяется скачком, а в одномерном — на $(\Delta E)^{-1/2}$. Хотя окружающий мир явно является трехмерным, двумерный и одномерный случаи представляют не только академический интерес. Так, плотности электронных состояний в полимерных цепочках или состояний вдоль дислокаций описываются одномерным случаем, а поверхностные состояния — двумерным.

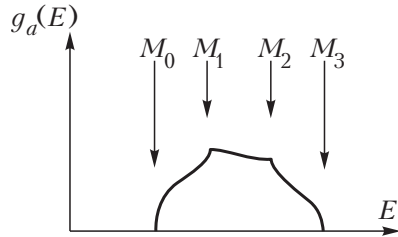


Рис. 14. Схематическое изображение особенностей Ван Хова в простой зоне, образованной невырожденным состоянием.

Плотность состояний в модели свободных электронов с эффективной массой m^* равна

$$g_{free}(E) = \frac{V(2m^*)^{3/2}}{4\pi^2 \hbar^3} \sqrt{E}. \quad (4.10)$$

С учетом вырождения по спину плотность состояний удваивается.

4.6 Заполнение одноэлектронных состояний в твердом теле

Итак, в результате какого-либо расчета получается картина одноэлектронных состояний и соответствующих им зон энергий в твердом теле. После этого необходимо заполнить полученные одноэлектронные состояния электронами твердого тела. При этом в принципе возможны две ситуации (рис. 15). Если верхняя зона, в которой еще находятся электроны, заполнена лишь частично, до уровня, имеющего энергию E_F (энергия Ферми), то такое вещество оказывается металлом, поскольку для перевода электрона в незанятое состояние, в котором он может двигаться по кристаллу, требуется очень малая энергия. В том случае, если верхняя валентная зона заполнена полностью и, чтобы достичь следующего электронного состояния, в котором электроны могут перемещаться по кристаллу, электрону необходимо приобрести заметную

энергию, вещество оказывается диэлектриком или полупроводником. Разница между этими двумя типами твердых тел в случае идеальных кристаллов в основном количественная — ширина запрещенной зоны у диэлектриков выше 3—4 эВ, у полупроводников — ниже этой цифры.

Поскольку, в соответствии с принципом Паули, в одном и том же электронном состоянии могут находиться не более двух электронов, имеющих противоположно направленные спины, кристаллы с нечетным числом электронов на элементарную ячейку всегда создают частично заполненные зоны и, соответственно, являются металлами. Обратное в общем случае не верно: из-за перекрытия нескольких зон многие кристаллы, имеющие четное число электронов на ячейку, тем не менее являются металлами.

Этой картине одноэлектронных состояний в твердом теле присуща определенная ограниченность. В гамильтониане \hat{H}_0 в процессе заполнения электронами состояний меняется вклад члена, описывающего электрон-электронное взаимодействие. В некоторых случаях это взаимодействие приводит к полной перестройке энергетического спектра, что в свою очередь может привести к переходу из металлического в диэлектрическое состояние (так называемые переходы Мотта (см., например, [63])). Поэтому одноэлектронные переходы обычно не учитывают этот вклад, называемый корреляционной энергией. Последовательное решение в рамках приближения Хартри-Фока, учитывающего корреляционные эффекты, обычно осуществить не удастся. Часто ограничиваются учетом некоторого среднего самосогласованного поля заполненных электронных состояний. Это приводит к разграничению описания одноэлектронных переходов (без учета влияния самосогласованного поля) и переходов в экситонные состояния.

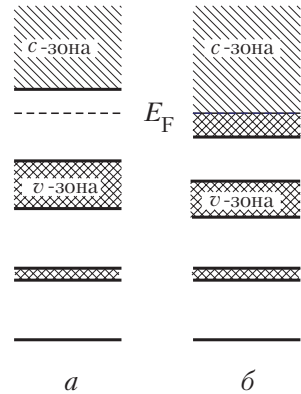


Рис. 15. Распределение энергетических зон в твердом теле в одноэлектронном приближении *a* - диэлектрик; *б* - металл. Двойной штриховкой отмечены заполненные зоны.

§ 5 Переходы под действием света в идеальном кристалле

5.1 Межзонные переходы

Рассмотрим переходы под действием кванта света из одной зоны в другую (так называемые межзонные переходы). При этом будем работать в рамках одноэлектронного приближения. Волновые функции электрона описываются функциями Блоха (4.4)

$$\psi_{\mathbf{k}\alpha} = V^{-1/2} e^{i\mathbf{k}\mathbf{r}} u_{\mathbf{k}\alpha}(\mathbf{r}),$$

которым соответствуют энергии $E_{\alpha}(\mathbf{k})$. Здесь α — номер энергетической зоны, \mathbf{k} — волновой вектор, определенный в зоне Бриллюэна. Матричный элемент перехода из начального состояния \mathbf{k}_i , α_i в конечное \mathbf{k}_f , α_f имеет вид

$$\langle f|\mathbf{e}\mathbf{r}|i\rangle = V^{-1} \int_V e^{-i\mathbf{k}_f\mathbf{r}} u_{\mathbf{k}_f\alpha_f}^*(\mathbf{r}) \mathbf{e}\mathbf{r} u_{\mathbf{k}_i\alpha_i}(\mathbf{r}) e^{i\mathbf{k}_i\mathbf{r}} d^3r.$$

Здесь интеграл берется по всему объему кристалла. Пользуясь периодичностью функций $u_{\mathbf{k}\alpha}$, можно свести этот интеграл к интегралу только по элементарной ячейке кристалла. Для этого определим вектор $\bar{\rho} = \mathbf{r} - \mathbf{n}$ — радиус-вектор какой-либо точки \mathbf{r} в элементарной ячейке относительно центра ячейки. При этом \mathbf{n} совпадает с одним из векторов трансляций (4.3). Матричный элемент разбивается на сумму по ячейкам объема v :

$$\begin{aligned} \langle f|\mathbf{e}\mathbf{r}|i\rangle &= \\ &= V^{-1} \sum_{\mathbf{n}} e^{-i(\mathbf{k}_f - \mathbf{k}_i)\mathbf{r}} \int_v u_{\mathbf{k}_f\alpha_f}^*(\bar{\rho}) \mathbf{e}\bar{\rho} u_{\mathbf{k}_i\alpha_i}(\bar{\rho}) e^{i(\mathbf{k}_f - \mathbf{k}_i)\bar{\rho}} d^3\rho. \end{aligned}$$

Член, пропорциональный $\mathbf{e}\mathbf{n}$, выпадает из-за ортогональности

функций $u_{\mathbf{k}\alpha}$. Сумма по \mathbf{n} может быть преобразована в сумму δ -функций, что приводит к закону сохранения импульса электрона:

$$\sum_{\mathbf{n}} e^{-i(\mathbf{k}_f - \mathbf{k}_i)\mathbf{r}} = \frac{8\pi^3}{V} \sum_{\mathbf{b}} \delta(\mathbf{k}_f - \mathbf{k}_i - \mathbf{b}).$$

Здесь $\mathbf{b} = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 + m_3 \mathbf{b}_3$, где \mathbf{b}_1 , \mathbf{b}_2 и \mathbf{b}_3 — базисные векторы обратной решетки. Принимая во внимание, что волновые векторы \mathbf{k}_i и \mathbf{k}_f определены в пределах первой зоны Бриллюэна, находим, что в сумме δ -функций остается только одна функция $\delta(\mathbf{k}_f - \mathbf{k}_i)$. Это справедливо только в использованном нами длинноволновом приближении. При более точном рассмотрении импульс электрона меняется на величину импульса поглощаемого фотона, что изменяет δ -функцию на $\delta(\mathbf{k}_f - \mathbf{k}_i - \mathbf{s} \omega n/c)$, где \mathbf{s} — единичный вектор в направлении распространения фотона. Тем самым в сумме из δ -функций остаются и члены, описывающие так называемые процессы переброса. Если \mathbf{k}_i выбран в пределах первой зоны Бриллюэна, то с учетом импульса фотона \mathbf{k}_f может выйти за пределы этой зоны. Для приведения этого вектора в первую зону необходимо вычесть из него вектор обратной решетки. Волновым вектором фотона невозможно пренебрегать (и тем самым, использовать длинноволновое приближение) в случае, если длина волны падающего излучения становится сравнимой или меньше постоянной решетки, то есть для достаточно жесткого рентгеновского излучения.

Мнимая часть диэлектрической проницаемости (3.8) принимает вид

$$\varepsilon_2(\omega) = 4\pi \sum_{\mathbf{k}, \alpha_f, \alpha_i} |\mathbf{e} \mathbf{d}_{fi}(\mathbf{k})|^2 [1 - f(E_f)] f(E_i) \delta(\hbar\omega + E_i(\mathbf{k}) - E_f(\mathbf{k})),$$

где

$$\mathbf{d}_{fi}(\mathbf{k}) = e v^{-1} \int_{\bar{v}} u_{\mathbf{k}_f \alpha_f}^*(\bar{\rho}) \bar{\rho} u_{\mathbf{k}_i \alpha_i}(\bar{\rho}) d^3 \rho.$$

Переходя от суммы по импульсам к интегралу с помощью соотношения $\sum_{\mathbf{k}} \dots = \frac{V}{8\pi^3} \int_{\Omega_B} \dots d^3 k$, получаем

$$\varepsilon_2(\omega) = \frac{1}{2\pi} \int d^3k \sum_{\alpha_f, \alpha_i} |\mathbf{e}d_{fi}(\mathbf{k})|^2 \times \quad (5.1)$$

$$\times [1 - f(E_f)] f(E_i) \delta(\hbar\omega + E_i(\mathbf{k}) - E_f(\mathbf{k})).$$

Эта формула описывает мнимую часть диэлектрической проницаемости для переходов электрона из одной зоны в другую. Интегрирование проводится по всей зоне Бриллюэна.

Для металлических кристаллов возможны переходы либо из частично заполненной энергетической зоны, либо в нее. При нулевой температуре интегрирование из-за наличия множителя $[1 - f_f] f_i$ производится либо внутри, либо вне поверхности Ферми металла.

Для диэлектриков возможны лишь межзонные переходы из заполненной валентной зоны ($f(E_i) = 1$) в свободную зону проводимости ($f(E_f) = 0$). В силу этого при низких температурах множитель $[1 - f_f] f_i$ можно опустить, ограничив суммирование лишь указанными зонами. Для диэлектриков это возможно при всех температурах, а для чистых полупроводников этот множитель описывает появление некоторого числа свободных равновесных электронов в зоне проводимости и, соответственно, такого же числа (для полупроводников без примесей) свободных состояний на потолке валентной зоны. Такие свободные состояния называются дырками. В дальнейшем в основном будут рассматриваться диэлектрики, поэтому мы будем считать нижнюю зону полностью заполненной, а верхнюю – свободной.

При поглощении кванта света происходят переходы электрона из состояния α_i в состояние α_f . Такое описание используется в случае, когда основное состояние рассматривается как состояние N электронов, заполняющих все валентные состояния. В этом случае полное число электронов во всех состояниях не меняется, а изменяются лишь их состояния. Далее состояния электронов в зоне проводимости мы будем отмечать индексом s , а в валентной зоне – v .

Однако часто удобнее рассматривать состояние твердого тела, используя понятия возбужденных состояний. Не возбужденное состояние, соответствующее заполненным валентным зонам, рассматривается как вакуум, в котором отсутствуют возбуждения. Поглощение фотона (или другой переход твердого тела в возбужденное состояние приводит в появлении пары: электрон e в зоне проводимости и дырка h в

валентной зоне. Далее при работе с одноэлектронными состояниями будем использовать индексы s и v , а для описания возбуждений — e и h .

Если начало энергии в картине одноэлектронных состояний выбрано на потолке валентной зоны, энергия такого электрона $E_e(\mathbf{k}_e)$ совпадает энергией одноэлектронного состояния в зоне проводимости $E_c(\mathbf{k}_c)$, в то время как $E_h(\mathbf{k}_h)$ равна по абсолютной величине и противоположна по знаку энергии соответствующего состояния в валентной зоне: $E_h(\mathbf{k}_h) = -E_v(-\mathbf{k}_v)$. Таким образом, энергии как электрона, так и дырки оказываются положительными величинами. Импульс дырки равен по абсолютной величине и противоположен по направлению импульсу соответствующего одноэлектронного состояния: $\mathbf{k}_h = -\mathbf{k}_v$ (если полный импульс системы из N электронов равен нулю, то импульс системы из $N - 1$ электронов противоположен импульсу удаленного электрона).

Законы сохранения энергии и импульса при поглощении длинноволнового фотона на языке возбужденных электронно-дырочных пар могут быть записаны как

$$\hbar\omega = E_c(\mathbf{k}_c) - E_v(\mathbf{k}_v) = E_e(\mathbf{k}_e) + E_h(\mathbf{k}_h) ,$$

$$0 = \mathbf{k}_c - \mathbf{k}_v = \mathbf{k}_e + \mathbf{k}_h .$$

Переходы этого типа называются прямыми межзональными переходами, поскольку на графике зависимости энергии одноэлектронных состояний E_α от квазиимпульса \mathbf{k} (рис. 16) эти переходы изображаются вертикальными линиями из-за сохранения импульса электрона.

Из-за наличия $\delta(\hbar\omega + E_i(\mathbf{k}) - E_f(\mathbf{k}))$ в формуле (5.1) переходы имеют пороговый характер, поскольку зона проводимости и валентная зона разделены по энергии. Функция

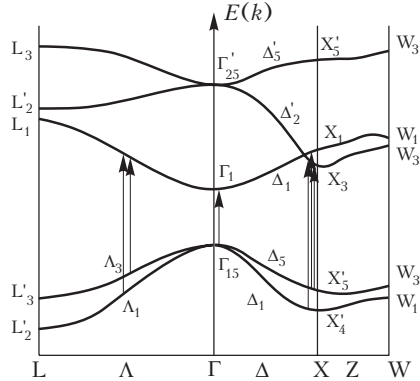


Рис. 16. Прямые межзональные переходы в диэлектрике со структурой NaCl.

$E_f(\mathbf{k}) - E_i(\mathbf{k})$ имеет минимум при определенном значении \mathbf{k}_{\min} , и переходы начинаются для данной пары ветвей при $\hbar\omega = E_{c,f}(\mathbf{k}_{\min}) - E_{v,i}(\mathbf{k}_{\min})$. Если экстремумы зон (минимум зоны проводимости и максимум валентной зоны) достигаются в различных точках зоны Бриллюэна, то значение пороговой энергии может быть больше, чем минимальное расстояние по энергии между валентной зоной и зоной проводимости. Квазиимпульс \mathbf{k}_{\min} определяется из условия

$$\left| \nabla_{\mathbf{k}} E_{c,f}(\mathbf{k}) - \nabla_{\mathbf{k}} E_{v,i}(\mathbf{k}) \right| = 0. \quad (5.2)$$

Последнее уравнение, вообще говоря, определяет все экстремумы \mathbf{k}_{ext} разности энергий зон. Это уравнение, во всяком случае, удовлетворяется в центре зоны Бриллюэна Γ и в точках высокой симметрии на ее границах, в которых в силу симметричных соображений зануляются производные функции $E_{\alpha}(\mathbf{k})$:

$$\nabla_{\mathbf{k}} E_{c,f}(\mathbf{k}) = \nabla_{\mathbf{k}} E_{v,i}(\mathbf{k}) = 0.$$

Уравнение (5.2) может также удовлетворяться в одной из точек вдоль осей симметрии третьего и четвертого порядков, вдоль которых обращаются в ноль компоненты градиентов, перпендикулярные осям.

Для исследования зависимости ε_2 от частоты перейдем к интегрированию по поверхностям постоянной разности энергий ветвей зоны проводимости и валентной зоны (это преобразование аналогично переходу от формулы (4.6) к (4.7)). В таком случае интеграл от δ -функции снимается, и мы приходим к равенству

$$\varepsilon_2(\omega) = \frac{1}{2\pi} \sum_{\alpha_f, \alpha_i} \oint_{E_f(\mathbf{k}) - E_i(\mathbf{k}) = \hbar\omega} \frac{|\mathbf{e}d_{fi}(\mathbf{k})|^2 dS_{\mathbf{k}}}{\left| \nabla_{\mathbf{k}} E_f(\mathbf{k}) - \nabla_{\mathbf{k}} E_i(\mathbf{k}) \right|}. \quad (5.3)$$

Если переходы между экстремальными точками ветвей α_i и α_f разрешены, то выражение $|\mathbf{e}d_{fi}(\mathbf{k}_{\text{ext}})|^2$ отлично от нуля и его можно вынести за знак интеграла при интегрировании вблизи экстремума. Вклад этих ветвей в ε_2 оказывается равным

$$\frac{1}{2\pi} \left| \mathbf{e} \mathbf{d}_{fi}(\mathbf{k}_{\text{ext}}) \right|^2 \oint_{E_f(\mathbf{k}) - E_i(\mathbf{k}) = \hbar\omega} \frac{|\mathbf{e} \mathbf{d}_{fi}(\mathbf{k})|^2 dS_{\mathbf{k}}}{\left| \nabla_{\mathbf{k}} E_f(\mathbf{k}) - \nabla_{\mathbf{k}} E_i(\mathbf{k}) \right|}.$$

Получающийся здесь интеграл по аналогии с плотностью состояний называется комбинированной межзонной плотностью состояний. В соответствии с этим комбинированная межзонная плотность состояний будет иметь не менее четырех особенностей Ван Хофа для каждой пары ветвей. Вблизи от экстремума можно разложить $E_f(\mathbf{k}) - E_i(\mathbf{k})$ в ряд аналогично разложению (4.8). Коэффициенты соответствующей квадратичной формы определяют тензор обратных приведенных масс μ_{ij}^{-1} . Этот тензор (в случае, если экстремумы обеих ветвей достигаются в одной и той же точке) равен разности тензоров обратных эффективных масс двух ветвей: $\mu_{ij}^{-1} = m_{c,ij}^{-1} - m_{v,ij}^{-1}$, или, на языке электронно-дырочных возбуждений, $\mu_{ij}^{-1} = m_{e,ij}^{-1} + m_{h,ij}^{-1}$. Таким образом приведенная межзонная масса оказывается меньше как массы электрона, так и массы дырки.

Для дипольно-разрешенных переходов вклады особенностей Ван Хофа в мнимую часть диэлектрической проницаемости ε_2 имеют различную форму в зависимости от числа отрицательных составляющих приведенного к главным осям тензора μ_{ij}^{-1} . Обозначения особых точек совпадают с обозначениями, приведенными в предыдущем параграфе:

$$\begin{aligned} M_0: \quad \Delta\varepsilon_2(\omega) &\sim \sqrt{\hbar\omega - E_{\text{ext}}} \theta(\hbar\omega - E_{\text{ext}}), \\ M_1: \quad \Delta\varepsilon_2(\omega) &\sim -\sqrt{E_{\text{ext}} - \hbar\omega} \theta(E_{\text{ext}} - \hbar\omega), \\ M_2: \quad \Delta\varepsilon_2(\omega) &\sim -\sqrt{\hbar\omega - E_{\text{ext}}} \theta(\hbar\omega - E_{\text{ext}}), \\ M_3: \quad \Delta\varepsilon_2(\omega) &\sim \sqrt{E_{\text{ext}} - \hbar\omega} \theta(E_{\text{ext}} - \hbar\omega), \end{aligned} \quad (5.4)$$

где $E_{\text{ext}} = E_{c,f}(\mathbf{k}_{\text{ext}}) - E_{v,i}(\mathbf{k}_{\text{ext}})$.

В ряде случаев по соображениям симметрии матричный элемент дипольного перехода $\mathbf{d}_{fi}(\mathbf{k}_{\text{ext}})$ обращается в ноль в точке экстремума. В таком случае необходимо учитывать следующую степень разложения матричного элемента перехода в ряд по степеням $\mathbf{k} - \mathbf{k}_{\text{ext}}$: $\mathbf{d}_{fi}(\mathbf{k}_{\text{ext}}) \sim \mathbf{k} - \mathbf{k}_{\text{ext}}$. В интеграле (5.3) квадрат $\mathbf{k} - \mathbf{k}_{\text{ext}}$ остается под знаком интеграла, и

сингулярность Ван Хова в этой точке становится значительно слабее: показатель степени в (5.4) меняется с $1/2$ на $3/2$. Обратим внимание на тот факт, что в случае дипольно-запрещенных переходов в твердом теле нет необходимости рассматривать магнитно-дипольные и квадрупольные переходы в точке экстремума, поскольку небольшой сдвиг по \mathbf{k} из этой точки обычно снимает запрет на переходы, и вклад в диэлектрическую проницаемость при энергиях немного выше пороговой энергии становится конечным и для дипольных переходов.

Поскольку спектр энергий имеет много ветвей как в валентной зоне, так и в зоне проводимости, в оптическом спектре поглощения должно наблюдаться большое число особенностей Ван-Хова. Для примера на рис. 17 изображены рассчитанный спектр ϵ_2 германия и экспериментально наблюдаемый спектр. На пунктирной кривой — теоретическом расчете переходов между нижними двумя зонами проводимости и валентной зоной — отчетливо проявляются особенности Ван-Хова. Аналогичные особенности наблюдаются и в экспериментально измеренном спектре. Некоторое различие в положении особенностей связано с приближениями, применявшимися при расчетах. Заметим, что острый пик в области $\sim 3,7$ эВ является результатом близкого расположения особенностей типа M_1 и M_2 . На рисунке видно качественное согласие между двумя типами кривых.

Одним из наиболее часто применяемым приближением при вычислении ϵ_2 является приближение постоянного (не зависящего от импульса) матричного элемента \mathbf{d}_{fi} . Это приближение никак не может быть обосновано, если импульс меняется в пределах всей зоны Бриллюэна, но используется из-за неточного определения функций $u_\alpha(\mathbf{r})$. В таком приближении ϵ_2 пропорционально сумме для всех пар ветвей от приведенной межзонной плотности состояний.

Сила

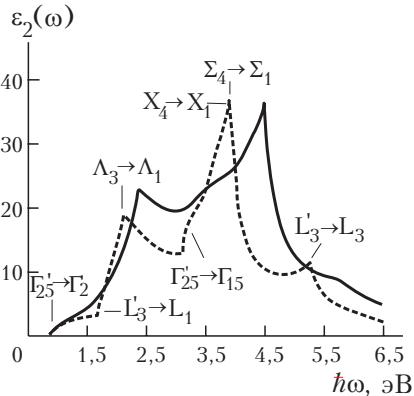


Рис. 17. Экспериментально измеренный спектр ϵ_2 для Ge и теоретический расчет ϵ_2 (по [64]).

осциллятора для соответствующих переходов не падает до нуля с постом энергии, и тем самым не может быть удовлетворено правило сумм, что означает неприменимость приближения постоянного матричного элемента в широкой области энергий. Учет зависимости матричного элемента перехода от волнового вектора может значительно изменить результат расчета оптических функций (на рис. 18 показаны результаты расчета [65]).

В металлах особенности оптических функций могут быть связаны не только с критическими точками комбинированной межзонной плотности состояний, но и с резким скачком электронной функции распределения $f(E)$ на поверхности Ферми. Однако такой эффект возможен лишь в том случае, когда поверхность Ферми достаточно близка к одной из поверхностей постоянной разности энергий S_ω . Тем самым при соответствующей энергии происходит резкий скачок числа возможных переходов, участвующих в поглощении света. Такое явление может возникать в металлах с плоскими валентными зонами.

Особое значение имеет особенность Ван Хофа в самом начале поглощения (на пороге поглощения), в связи с тем, что при этом поглощение должно резко обрываться в сторону меньших энергий по закону $(\hbar\omega - E_g)^{1/2}$ для разрешенных переходов и $(\hbar\omega - E_g)^{3/2}$ для запрещенных. Однако здесь уже начинаются трудности применения одноэлектронной модели без учета корреляционных эффектов. Если для ряда полупроводников, например Ge, поведение экспериментально

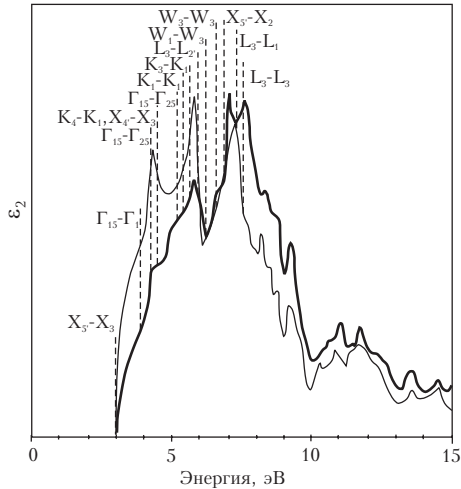


Рис. 18. Мнимая часть диэлектрической проницаемости CaS с учетом (толстая линия) и без учета (тонкая линия) зависимости матричного элемента перехода от волнового вектора [65].

измеренного ε_2 на краю качественно (с точностью до относительно малых эффектов, связанных с непрямыми переходами) совпадает с теоретическими расчетами, то для многих широкозонных кристаллов (кристаллов с большей шириной запрещенной зоны, $E_g > 5$ эВ) наблюдаются резкие отличия между экспериментом и простой теорией. Например, в экспериментально наблюдаемых спектрах поглощения щелочно-галогенидных кристаллов имеются резкие максимумы на краю фундаментального поглощения, которые не могут быть объяснены в рамках простых одноэлектронных представлений без учета взаимодействия между электронами.

5.2 Пространственная дисперсия межзонной диэлектрической проницаемости

Как уже было сказано, в спектроскопии твердых тел важную роль играет не только частотная, но и пространственная дисперсия оптических функций, в частности, $\varepsilon_2(\omega, \mathbf{k})$. Учет этого явления важен не только при описании взаимодействия твердого тела с поперечными фотонами – квантами света, но и при рассмотрении взаимодействия с продольными фотонами кулоновского поля. Поглощение таких фотонов с энергией $\hbar\omega$ и импульсом $\hbar\mathbf{k}$ приводит к соответствующим потерям энергии и импульса пролетающей заряженной частицы. Выражение для $\varepsilon_2(\omega, \mathbf{k})$ с учетом зависимости от волнового вектора \mathbf{k} аналогично (5.1) с единственным различием, что волновой вектор продольного фотона войдет в закон сохранения импульса:

$$\mathbf{k} = \mathbf{k}_c - \mathbf{k}_v = \mathbf{k}_e + \mathbf{k}_h.$$

Естественно, что это изменение изменяет порог поглощения, поскольку \mathbf{k} войдет в уравнения для определения экстремума. Матричный элемент \mathbf{d}_{fi} также зависит от \mathbf{k} . Но даже в этом случае в $\varepsilon_2(\omega, \mathbf{k})$ проявляются особенности Ван Хофа по ω , положение которых зависит от \mathbf{k} . В частности, если рассматривать только область вблизи порога, то для разрешенных переходов будет наблюдаться корневая особенность. В простейшем случае дисперсионные кривые для электронов и дырок могут быть записаны в параболическом приближении:

$$E_e(\mathbf{k}) = E_g + \frac{\hbar^2(\mathbf{k} - \mathbf{k}_0)^2}{2m_e}, \quad E_h(\mathbf{k}) = \frac{\hbar^2\mathbf{k}^2}{2m_h}.$$

(экстремум валентной зоны предполагается находящимся в точке $\mathbf{k} = 0$, а экстремум зоны проводимости — в $\mathbf{k} = \mathbf{k}_0$). Если $\mathbf{k}_0 \neq 0$, величина E_g соответствует непрямой запрещенной зоне. В таком случае $\varepsilon_2(\omega, \mathbf{k})$ принимает форму

$$\varepsilon_2(\omega, \mathbf{k}) \sim \sqrt{\hbar\omega - E_g - \frac{\hbar^2(\mathbf{k} - \mathbf{k}_0)^2}{2(m_e + m_h)}}. \quad (5.5)$$

Порог по $\hbar\omega$ для поглощения продольного фотона зависит от \mathbf{k} :

$$E_{\min}(\mathbf{k}) = E_g + \frac{\hbar^2(\mathbf{k} - \mathbf{k}_0)^2}{2(m_e + m_h)}.$$

В соответствии с этой формулой изменяются минимальные потери энергии быстрой частицы, передающей твердому телу импульс $\hbar\mathbf{k}$.

Минимальная передаваемая энергия E_{\min} растет с ростом \mathbf{k} , если положения экстремумов двух зон совпадают ($\mathbf{k}_0 = 0$). Но в случае $\mathbf{k}_0 \neq 0$ минимальная передаваемая твердому телу энергия при непрямых переходах (всегда равная $E_g^i = E_g$) меньше, чем порог прямого поглощения поперечного фотона (в рассматриваемом случае $E_g^d = E_g + \frac{\hbar^2 k_0^2}{2(m_e + m_h)}$).

Таким образом, поперечные фотоны с энергией между E_g и E_g^d могут поглощаться в твердом теле только в процессах с участием третьих частиц, например, дефектов или фононов (более подробно это рассмотрено в третьей главе).

Если экстремум зоны достигается не в Γ -точке, то в силу симметричных соображений таких экстремальных точек должно быть несколько. Тем самым выражение (5.5) необходимо просуммировать по всем эквивалентным экстремумам, и симметрия функции ε_2 будет полностью определяться симметрией обратной решетки кристалла.

Поскольку зависимость ε_2 от \mathbf{k} в кристалле не является изотропной, вычисление полной зависимости оптических функций от переменных ω и \mathbf{k} не проводится из-за громоздкости. Чаще всего ограничиваются изучением зависимости ε_2 от \mathbf{k} при нулевой частоте, что важно для исследования статического экранирования в кристаллах. Но для правильного

понимания явлений, связанных с неупругими электрон-электронными взаимодействиями необходимо учитывать, хотя бы на качественном уровне, зависимость ϵ_2 как от ω , так и от \mathbf{k} .

5.3 Эффект Франца-Келдыша

Если к кристаллу приложено постоянное электрическое поле \mathcal{E} , то поглощение вблизи порога изменяется. Этот эффект известен как эффект Франца-Келдыша. Электрическое поле приводит к наклону разрешенных зон в координатном пространстве (рис. 19). Если поле \mathcal{E} приложено вдоль оси x , все энергии становятся зависящими от координат из-за потенциальной энергии $-e\mathcal{E}x$, которую приобретает электрон в таком поле. В таком случае становится возможным туннельный переход электрона из валентной зоны в зону проводимости.

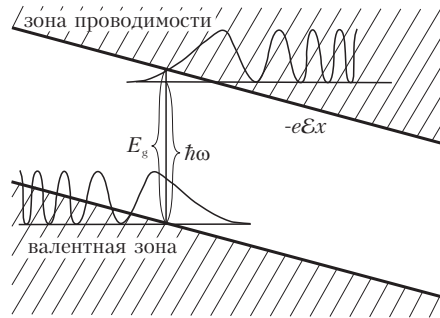


Рис. 19. Разрешенные зоны и схематическое изображение волновых функций в кристалле, к которому приложено электрическое поле \mathcal{E} .

Пространственная ширина щели, разделяющей эти состояния, обратно пропорциональна приложенному полю \mathcal{E} .

Волновые функции электронов в зонах уже не описываются блоховскими функциями. Они остаются периодически по осям, перпендикулярным оси x (если поле направлено по оси симметрии). Их зависимость от координаты x описывается довольно громоздкими функциями (типичное поведение $|\psi|^2$ показано на рис. 19), и поэтому обычно переходят к импульсному представлению (см., например, [4]). Матричный элемент перехода между валентной зоной и зоной проводимости оказывается пропорциональным функции Эйри. Анализ этих решений довольно громоздок, и поэтому здесь мы приведем только конечные результаты.

Мнимая часть диэлектрической проницаемости ниже порога поглощения E_g в присутствии электрического поля

оказывается пропорциональной

$$\varepsilon_2(\omega, \mathcal{E}) \sim \exp\left[-\gamma \frac{(E_g - \hbar\omega)^{3/2}}{\mathcal{E}}\right], \quad \hbar\omega < E_g, \quad (5.6)$$

где $\gamma = 4\sqrt{2\mu}/3e\hbar$. Выше порога поглощения коэффициент поглощения осциллирует с периодом, уменьшающимся с падением электрического поля \mathcal{E} . Эти осцилляции связаны с отражением электронов и дырок от барьера, создаваемого электрическим полем. Стоит отметить чисто квантовую природу эффекта Франца-Келдыша, поскольку в пределе $\hbar \rightarrow 0$ поглощение исчезает.

Отметим еще связанное с эффектом Франца-Келдыша явление, заключающееся во внутренней эмиссии электронов в электрическом поле (эффект Зинера) в отсутствие высокочастотного электромагнитного поля. Вероятность перехода электрона из одной зоны в другую пропорциональна экспоненциальному множителю (5.6), в котором полагается $\omega = 0$.

5.4 Поглощение фотонов большой энергии

Поглощение фотонов достаточно большой энергии, соответствующей рентгеновскому диапазону спектра, приводит к переходам электрона из глубоких остовных зон в зону проводимости. Протяженность волновых функций исходного состояния и конечного состояния существенно отличаются друг от друга. Это связано с тем, что характерные размеры остовных состояний порядка или меньше боровского радиуса для атома водорода. При таких обстоятельствах нет оснований для предположений о постоянстве матричных элементов перехода.

В нулевом приближении волновая функция электрона в зоне проводимости с энергией в десятки и сотни электрон-вольт может быть рассмотрена в приближении свободных электронов, и при этом она имеет вид

$$\psi_{\mathbf{k}} = V^{-1/2} e^{i\mathbf{k}\mathbf{r}}.$$

Остовные состояния могут быть рассмотрены в противоположном приближении, то есть в приближении сильно связанных электронов, в котором волновая функция строится в виде суперпозиции остовных функций. В таком случае коэффициент поглощения из остовного состояния в нулевом

приближении пропорционален вероятности фотоионизации изолированного атома. Значение матричного элемента дипольного момента определяется асимптотическим поведением функции остоного состояния на больших расстояниях. Для остоного состояния эта асимптотика является экспонентой, показатель которой растет с ростом энергии ионизации соответствующего уровня E_I . Тем самым матричный элемент дипольного момента сильно зависит от волнового вектора конечного состояния $\mathbf{p} = \hbar\mathbf{k}$, определяемого энергией падающего фотона $p = \sqrt{2m(\hbar\omega - E_I)}$:

$$\langle \mathbf{k} | \mathbf{er} | i \rangle \sim \frac{e\mathbf{p}}{(p^2 + mE_I)^2} \quad (5.7)$$

(более подробно вывод этой формулы рассматривается в четвертой части, где обсуждается фотоионизация дефектных состояний в твердом теле). Тем самым коэффициент поглощения в широкой области энергий в одноэлектронном приближении пропорционален

$$\varepsilon_2(\omega) \sim \omega^{-7/2}, \quad \hbar\omega \gg E_I. \quad (5.8)$$

Для энергий фотона вблизи порога поглощения с остоного состояния формула (5.8) становится неприменимой по ряду причин. При этом она не применима и в изолированных атомах из-за многоэлектронных эффектов, в результате которых изменяется матричный элемент (5.7). Эти изменения можно подразделить на несколько групп.

1. Переходы из остоных состояний, которые характеризуются определенным значением орбитального квантового числа l , происходят не в состояния плоских волн, а в суперпозицию плоских волн, построенную из функций с одинаковым абсолютным значением волнового вектора \mathbf{k} и являющихся сферическими волнами с орбитальным моментом l' . Угловая часть этой суперпозиции описывается сферической функцией $Y_{l'm}(\theta, \varphi)$. В соответствии с правилами отбора матричный элемент дипольного перехода не равен нулю только для переходов $l \rightarrow l' = l \pm 1$. Переходы в другие состояния в дипольном приближении отсутствуют, и появляются только при учете более высоких мультиполей. Переходы в состояния с большим орбитальным квантовым числом $l \rightarrow l' = l + 1$ обычно имеют интенсивность на

один-два порядка выше, чем переходы с уменьшением $l \rightarrow l' = l - 1$. Более подробное рассмотрение этих вопросов приводится в литературе по атомной спектроскопии (см., например, [66–68]).

2. Максимум радиальной части волновой функции для состояний с $l \neq 0$ сдвигается из точки $\mathbf{r} = 0$ из-за наличия центробежного потенциала $l(l+1)\hbar^2/2mr^2$. Тем самым матричный элемент перехода из основного состояния в состояние в зоне проводимости приобретает максимальное значение не на пороге ионизации, а при больших энергиях, для которых радиальная часть волновой функции свободного электрона также имеет максимум, сдвинутый из начала координат. Тем самым полоса поглощения может иметь максимум, существенно сдвинутый от порога ионизации E_I , и этот сдвиг растет с ростом l .
3. Структура волновой функции начального состояния проявляется в зависимости формы линии поглощения от главного квантового числа n основного уровня. Это связано с тем, что число нулей в радиальной части волновой функции растет с ростом n . Матричные элементы перехода из таких состояний в состояния непрерывного спектра имеют немонотонную зависимость от энергии, что приводит к возникновению структуры в полосах поглощения.
4. Сдвиг полосы поглощения и осцилляции в спектре поглощения возникают при энергиях порядка единиц и десятков электрон-вольт над порогом. Все приведенные выше эффекты (1–3) имеют атомную природу. В этой же области спектра проявляются и эффекты, связанные со структурой твердого тела. В частности, нарушается приближение свободных электронов, использовавшееся для описания конечного состояния. В таком расчете необходимо использовать реальные блоховские функции и учитывать многоэлектронные эффекты. Эта область поглощения используется в методе XANES (X-ray Absorption Near Edge Structure – Структура рентгеновского поглощения вблизи края) для исследования структуры вещества путем расчета электронных состояний в зоне проводимости и в основной зоне.

Абсолютные значения максимума поглощения могут быть грубо оценены с использованием правил сумм (1.29). Если предположить, что поглощение начинается при энергиях, приблизительно равных E_I , и протяженность области

сильного поглощения составляет ΔE (ΔE меньше или порядка E_I), то среднее значение $\bar{\varepsilon}_2$ в этой области равно

$$\bar{\varepsilon}_2 \approx \frac{2\pi^2 e^2 \hbar^2 N_c}{m E_I \Delta E a^3} = \frac{\pi \hbar^2 \omega_{pl,c}^2}{2 E_I \Delta E},$$

где N_c — число электронов на соответствующем остоном уровне на одну элементарную ячейку, a — постоянная решетки, а $\omega_{pl,c}$ — плазменная частота, соответствующая данной группе остовных электронов. Поскольку в области рентгеновского поглощения $n \approx 1$, то из (1.11) и из определения коэффициента поглощения α (1.17) следует, что среднее значение коэффициента поглощения равно $\bar{\alpha} = \omega \bar{\varepsilon}_2 / c$. Поскольку $\omega \approx E_I / \hbar$, то получаем

$$\bar{\alpha} \approx \frac{\hbar \omega_{pl,c}^2}{c \Delta E}.$$

Оценим $\bar{\alpha}$ для вещества с $N_c = 2$ и постоянной решетки $a \approx 0,25$ нм. В таком случае $\hbar \omega_{pl,c} \approx 13$ эВ, откуда $\bar{\alpha} \approx 10^7 / \Delta E$ см⁻¹ (ΔE измеряется в электрон-вольтах). Если область фундаментального поглощения имеет ширину $\Delta E \approx 10$ эВ, то $\bar{\alpha}$ в этой области порядка 10^6 см⁻¹, а $\bar{\varepsilon}_2 \approx 1$. Для рентгеновского поглощения с энергией порога $E_I \approx 10$ КэВ с шириной $\Delta E \approx 1$ КэВ получаем $\bar{\alpha} \approx 10^4$ см⁻¹, а $\bar{\varepsilon}_2 \approx 10^{-5}$.

5.5 EXAFS

В интервале между областью XANES и областью, в которой коэффициент поглощения описывается формулой (5.8), имеется широкая область поглощения. Метод EXAFS (Extended X-ray Absorption Fine Structure – Протяженная тонкая структура рентгеновского поглощения) успешно применяется для исследования структуры вещества. В спектральной области, которая начинается при энергиях порядка 50 эВ над порогом E_I и растягивается на сотни электрон-вольт, наблюдаются осцилляции коэффициента поглощения, отсутствующие в атомных спектрах фотоионизации. Эти осцилляции связаны также с непостоянством матричных элементов перехода из остовных состояний в состоянии непрерывного спектра. Такие осцилляции могут быть описаны с учетом поправок первого порядка к волновым функциям почти свободных

электронов. Волновая функция электрона с энергией в десятки и сотни электронвольт в твердом теле отличается от волновой функции в свободном пространстве из-за суперпозиции с плоскими волнами, отражающимися от различных атомов [69].

Рассмотрим поглощение на основном состоянии конкретного атома A , расположенного в начале координат. Поскольку волновая функция начального основного состояния сильно локализована в пространстве вокруг атома A , матричный элемент дипольного момента перехода будет определяться не всей волновой функцией конечного состояния, а только областью этой функции вблизи атома A . Удобно строить волновую функцию электрона в таком состоянии, основываясь не на плоской волне, а на сферической волне, распространяющейся от центрального атома, на котором происходит поглощение, и на волнах, отражающихся от соседних атомов. Волновая функция сферической волны имеет вид $h_1^{(1)}(kr)Y_{lm}(\Omega_{\mathbf{r}})$, где $h_1^{(1)}$ — сферическая функция Ханкеля первого рода, $\Omega_{\mathbf{r}} = \{\theta_{\mathbf{r}}, \varphi_{\mathbf{r}}\}$ — угловые координаты, соответствующие вектору \mathbf{r} . Волновой вектор k определяется энергией электрона $\hbar\omega - E_I$:

$$k = \sqrt{2m(\hbar\omega - E_I)}/\hbar = \sqrt{(\hbar\omega - E_I)/Ry}/a_B ,$$

От соседнего атома с номером j , расположенного в точке \mathbf{R}_j , отражается сферическая волна, возбужденная первичной волной:

$$h_1^{(1)}(kR_j)Y_{lm}(\Omega_{\mathbf{R}_j})S(\theta, k)e^{ik|\mathbf{r}-\mathbf{R}_j|}|\mathbf{r}-\mathbf{R}_j|^{-1} , \quad (5.9)$$

где $S(\theta, k)$ — амплитуда волны с волновым вектором k , отраженной от j -го атома в направлении θ по отношению к направлению волны \mathbf{k} , падающей на данный атом. Поскольку для вычисления матричного элемента перехода необходимо знать волновую функцию только вблизи начала координат, возможно вычислить волновую функцию, раскладывая (5.9) в ряд около $\mathbf{r} = 0$ и используя асимптотическое представление для радиальной сферической функции:

$$h_1^{(1)}(kr)Y_{lm}(\Omega_{\mathbf{r}}) + \sum_j e^{i[2kR_j - (l+1)\pi/2]} \frac{Y_{lm}(\Omega_{\mathbf{R}_j})S(\pi, k)}{kR_j^2} e^{-\frac{ik\mathbf{R}_j\mathbf{r}}{R_j}} .$$

Если предположить, что начальное остовное состояние является K -состоянием (то есть $l = 0$), то квадрат модуля матричного элемента перехода, рассчитанный для чисто атомного перехода, умножается на множитель

$$1 + \frac{1}{k} \sum_j \frac{N_j}{R_j^2} |S_j(\pi, k)| \sin(2kR_j + \varphi_j(k)), \quad (5.10)$$

где суммирование производится по координационной сфере, содержащей N_j атомов, расположенных на расстоянии R_j от центрального атома. Здесь введены амплитуда и фаза рассеяния электрона с волновым вектором k назад атомом с номером j : $S_j(\pi, k) = |S_j(\pi, k)| e^{i\varphi_j(k)}$. Множитель (5.10) незначительно модифицируется для L и M начальных состояний атома A .

Таким образом, мнимая часть диэлектрической и коэффициент поглощения домножаются на функцию (5.10), что приводит к осцилляциям коэффициента поглощения. Период осцилляций в k -пространстве равен $\Delta k = \pi/R_j$ и численно порядка 1 нм^{-1} , что эквивалентно осцилляциям по энергии с периодом порядка $50 - 150 \text{ эВ}$ в зависимости от энергии электрона $\hbar\omega - E_I$. Если энергия фотона превышает порог на $500 - 1000 \text{ эВ}$, осцилляции затухают. Это связано с неупругим рассеянием электронов в зоне проводимости (обычно это учитывается введением в изменяющуюся часть (5.10) дополнительного множителя $e^{-2R_j/l(k)}$, где $l(k)$ — средняя длина пробега электрона с волновым вектором k ; подробности приведены в четвертой главе учебного пособия). Кроме того, необходимо учитывать колебания атомов вблизи положений равновесия и возможный структурный беспорядок. Эти две причины приводят к распределению межатомных расстояний R_j вокруг положения равновесия с функцией распределения

$$\sim \exp[-(R_j - R_j^0)^2 / 2\sigma_j],$$

где дисперсия σ_j равна сумме $\sigma_j = \sigma_j^T + \sigma_j^S$ дисперсии σ_j^T , связанной с колебаниями атомов, и σ_j^S , определяемой структурным беспорядком. Дисперсия σ_j^T зависит от температуры и уменьшается с ее падением, до величины σ_j^0 , определяемой нулевыми колебаниями. Дисперсия σ_j^S от температуры не

зависит и содержит информацию о неупорядоченности системы (и полезна для исследования стекол и других систем с нарушениями ближнего порядка). Такого рода распределения приводят к возникновению фактора Дебая-Уоллера $\exp(-2\sigma_j^2 k^2)$.

Дополнительные усложнения в формулу (5.10) вносит учет начального состояния — дырки на атоме A , что приводит к появлению дополнительного сдвига фаз в аргументах синусов. Более детальное рассмотрение формирования EXAFS-сигнала учитывает многократное рассеяние электроном системой окружающих атомов, однако при этом сложность расчетов спектров поднимается до уровня расчетов зонных структур. Многократное рассеяние необходимо учитывать в переходной области между областями EXAFS и XANES.

С помощью метода EXAFS можно оценивать расстояния между атомом A и окружающими его атомами. Для этого проводят преобразование Фурье по переменной k добавки к атомному спектру поглощения, умноженной на k . Положения пиков этого преобразования соответствуют расстояниям до различных координационных сфер, а интенсивность этих пиков пропорциональна числу атомов в данной сфере. В связи с такой простой интерпретацией результатов метод EXAFS как метод структурного анализа получил широкое распространение, особенно с применением синхротронного излучения с непрерывным рентгеновским спектром.

Как это следует из предыдущего анализа, метод EXAFS дает информацию о зависимости волновых функций от энергии высоко в зоне проводимости и о ее локальной структуре вблизи возбужденного атома. Исследуя поглощение с основных состояний других атомов, можно получить информацию о структуре той же функции в окрестностях других атомов, которые находятся и в другом окружении.

Поскольку поглощение сильно влияет на все вторичные процессы, был разработан ряд методов (измерение флюоресценции, фотопроводимости и фотоакустического эффекта) по альтернативным способам получения EXAFS-информации. Пример люминесцентного метода EXAFS приведен в четвертой главе учебного пособия.

§ 6 Экситонные эффекты в твердых телах

6.1 Кулоновское взаимодействие электрона и дырки

При предыдущем рассмотрении электронных эффектов в твердых телах член в гамильтониане, описывающий взаимодействие между электронами, учитывался в лучшем случае в усредненном по основному состоянию виде. При этом его вклад не изменялся в зависимости от электронной конфигурации системы. Для описания основного состояния системы такое приближение может быть оправдано. В то же время возможность пренебрежения межэлектронным взаимодействием при описании возбужденного состояния системы и особенно переходов между состояниями представляется куда более сомнительным.

При переходе с изменением состояния одного электрона изменение межэлектронного взаимодействия связано с возникновением электрона в зоне проводимости и дырки в валентной зоне. Вернемся к общему виду гамильтониана для кристалла с недеформируемой решеткой (4.2) и запишем его, выделяя в явном виде члены с межэлектронным взаимодействием:

$$\hat{\mathcal{H}}_0 = \sum_i \hat{h}(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (6.1)$$

Здесь в член $\hat{h}(\mathbf{r})$ включены энергия электрона и энергия его взаимодействия с ионными остовами. Будем для простоты предполагать, что система представляет собой диэлектрик, имеющий только две невырожденные зоны: валентную зону v и зону проводимости c . В качестве функций начального приближения рассмотрим блоховские одноэлектронные функции $\psi_{\mathbf{k}\alpha}(\mathbf{r})$ (4.4).

Для описания многоэлектронной системы удобно использовать операторы рождения и уничтожения частиц в

представлении вторичного квантования [49]. Будем обозначать оператор рождения электрона в состоянии с волновой функцией $\psi_{\mathbf{k}\alpha}$ как $a_{\mathbf{k}\alpha}^+$, а оператор уничтожения электрона в таком состоянии — как $a_{\mathbf{k}\alpha}$. Индекс α может принимать два значения: v и c в соответствии с предположением о невырожденности зон. Основным состоянием твердого тела является состояние с заполненной валентной зоной:

$$|g\rangle = \prod_{\mathbf{k}} a_{\mathbf{k}v}^+ |0\rangle.$$

Для диэлектриков произведение в этой формуле захватывает всю зону Бриллюэна, а для металлов распространяется только на область внутри поверхности Ферми. По причинам, рассмотренным ниже, в данном параграфе речь пойдет только о диэлектриках (или полупроводниках).

Если один из электронов из валентной зоны переходит в зону проводимости, то образуется состояние с полным импульсом $\hbar\mathbf{K}$ и импульсом электрона $\hbar\mathbf{k}$:

$$|\mathbf{kK}\rangle = a_{\mathbf{k}c}^+ a_{\mathbf{k}-\mathbf{K},v} |g\rangle.$$

При этом в валентной зоне образуется дырка с импульсом $\hbar(\mathbf{K} - \mathbf{k})$ (как указывалось в предыдущем параграфе, импульс дырки равен по модулю импульсу электрона, изъятых из валентной зоны, и направлен в противоположную сторону). Такое состояние можно описывать либо используя операторы рождения электрона в зоне проводимости и уничтожения электрона в валентной зоне (то есть на языке одноэлектронных состояний, используя индексы c и v), либо (на языке возбужденных состояний с индексами e и h) пользуясь операторами рождения электрона $a_{\mathbf{k}e}^+ = a_{\mathbf{k}c}^+$ и дырки $a_{\mathbf{k}h}^+ = a_{-\mathbf{k}v}$. Операторы уничтожения, соответствующие таким операторам рождения, имеют вид $a_{\mathbf{k}e} = a_{\mathbf{k}c}$ и $a_{\mathbf{k}h} = a_{-\mathbf{k}v}^+$.

В качестве основного приближения при описании взаимодействия электронов используется приближение Хартри — Фока [49]. Одноэлектронная энергия в этом приближении зависит только от волнового вектора \mathbf{k} блоховской функции:

$$E_{\alpha}^{\text{HF}}(\mathbf{k}) = \langle \alpha \mathbf{k} | \hat{h} | \alpha \mathbf{k} \rangle + \\ + \sum_{\alpha' \mathbf{k}'} \left(\left\langle \alpha \mathbf{k}, \alpha' \mathbf{k}' \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \alpha' \mathbf{k}', \alpha \mathbf{k} \right\rangle - \left\langle \alpha' \mathbf{k}', \alpha \mathbf{k} \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \alpha' \mathbf{k}', \alpha \mathbf{k} \right\rangle \right).$$

Здесь введено обозначение

$$\left\langle \alpha_1 \mathbf{k}_1, \alpha_2 \mathbf{k}_2 \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \alpha_3 \mathbf{k}_3, \alpha_4 \mathbf{k}_4 \right\rangle = \\ = \int d^3 r_1 d^3 r_2 \psi_{\alpha_1 \mathbf{k}_1}^*(\mathbf{r}_1) \psi_{\alpha_2 \mathbf{k}_2}^*(\mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_{\alpha_3 \mathbf{k}_3}(\mathbf{r}_2) \psi_{\alpha_4 \mathbf{k}_4}(\mathbf{r}_2).$$

В состояниях $|\mathbf{k}\mathbf{K}\rangle$ матричный элемент гамильтониана (6.1) диагонален по полному импульсу $\hbar\mathbf{K}$ и имеет вид

$$\langle \mathbf{k}'\mathbf{K} | \hat{\mathcal{H}}_0 | \mathbf{k}\mathbf{K} \rangle = \delta_{\mathbf{k}\mathbf{k}'} [E_c^{\text{HF}}(\mathbf{k}) - E_v^{\text{HF}}(\mathbf{k} - \mathbf{K})] - \\ - \left\langle c\mathbf{k}', v(\mathbf{k} - \mathbf{K}) \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| v(\mathbf{k}' - \mathbf{K}), c\mathbf{k} \right\rangle + \\ + \left\langle v(\mathbf{k} - \mathbf{K}), c\mathbf{k}' \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| v(\mathbf{k}' - \mathbf{K}), c\mathbf{k} \right\rangle.$$

Видно, что матричные элементы $\langle \mathbf{k}'\mathbf{K} | \hat{\mathcal{H}}_0 | \mathbf{k}\mathbf{K} \rangle$ не являются диагональными по импульсу электрона $\hbar\mathbf{k}$. Таким образом, при учете межэлектронного взаимодействия состояние с определенным импульсом электрона $\hbar\mathbf{k}$ не является собственным состоянием $\hat{\mathcal{H}}_0$. Для того чтобы получить собственную функцию гамильтониана $\hat{\mathcal{H}}_0$, необходимо составить линейную комбинацию из функций, имеющих одинаковый полный импульс $\hbar\mathbf{K}$:

$$|\lambda\mathbf{K}\rangle = \sum_{\mathbf{k}} f_{\lambda\mathbf{K}}(\mathbf{k}) |\mathbf{k}\mathbf{K}\rangle,$$

и, составив $\sum_{\mathbf{k}}$ секулярное уравнение, определить коэффициенты $f_{\lambda\mathbf{K}}(\mathbf{k})$, соответствующие собственному значению $E_{\lambda\mathbf{K}}$.

Прежде чем проделать это, введем альтернативное описание взаимодействия электрона с дыркой, которое явным образом учитывает расстояние между электроном и дыркой.

Перейдем от функций Блоха к функциям Ванье [70], связанным с первыми преобразованиями Фурье:

$$\Psi_{\alpha}(\mathbf{r} - \mathbf{n}) = N^{-1/2} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{n}} \psi_{\alpha\mathbf{k}}(\mathbf{r}) .$$

Здесь N — число атомов в основной области кристалла, на которую накладываются периодические условия Борна-Кармана. Функции Ванье $\Psi_{\alpha}(\mathbf{r} - \mathbf{n})$ имеют максимум в ячейке с дискретным радиус-вектором \mathbf{n} (4.3). В случае применимости приближения сильно связанных электронов в нулевом приближении функции Ванье просто совпадают с волновой функцией электрона для атома, находящегося в ячейке \mathbf{n} . Такое состояние включает в себя суммирование по всем состояниям с различными волновыми векторами и не является собственным состоянием гамильтониана даже при пренебрежении межэлектронным взаимодействием, но здесь его удобно использовать, поскольку для него электрон находится в основном около \mathbf{n} -ной ячейки. Оператор рождения электрона в этом состоянии равен

$$a_{\alpha\mathbf{n}}^{+} = N^{-1/2} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{n}} a_{\alpha\mathbf{k}}^{+} .$$

Вместо состояний $|\mathbf{k}\mathbf{K}\rangle$ для решения задачи на собственные значения можно использовать состояния $|\mathbf{I}\mathbf{K}\rangle$, построенные на функциях Ванье, которые соответствуют тому же самому полному моменту $\hbar\mathbf{K}$ и описывают электрон и дырку, находящиеся на расстоянии \mathbf{I} друг от друга:

$$|\mathbf{I}\mathbf{K}\rangle = N^{-1/2} \sum_{\mathbf{m}} e^{i\mathbf{K}\mathbf{m}} a_{c, \mathbf{m}+1}^{+} a_{v\mathbf{m}} |g\rangle .$$

Тогда состояние $|\lambda\mathbf{K}\rangle$ с определенным значением энергии $E_{\lambda\mathbf{K}}$ может быть записано через сумму таких функций:

$$|\lambda\mathbf{K}\rangle = \sum_{\mathbf{I}} F_{\lambda\mathbf{K}}(\mathbf{I}) |\mathbf{I}\mathbf{K}\rangle ,$$

где функция $F_{\lambda\mathbf{K}}(\mathbf{I})$ связана с введенными выше коэффициентами $f_{\lambda\mathbf{K}}(\mathbf{k})$ с помощью Фурье-преобразования:

$$F_{\lambda\mathbf{K}}(\mathbf{I}) = N^{-1/2} \sum_{\mathbf{k}} e^{i\mathbf{k}\mathbf{I}} f_{\lambda\mathbf{K}}(\mathbf{k}) .$$

По построению функция $f_{\lambda\mathbf{K}}(\mathbf{k})$ связана с вероятностью найти

электрон и дырку, имеющие относительный импульс $\hbar\mathbf{k}$ в состоянии $|\lambda\mathbf{K}\rangle$, в то время как функция $F_{\lambda\mathbf{K}}(\mathbf{l})$ — с вероятностью найти в том же состоянии электрон на расстоянии \mathbf{l} от дырки.

Состояние $|\lambda\mathbf{K}\rangle$, будучи собственным состоянием гамильтониана \hat{H}_0 , и учитывающее межэлектронное взаимодействие, называется экситонным состоянием. Секулярное уравнение для нахождения собственного значения гамильтониана имеет вид (см., например, [67, 71]):

$$\sum_{\mathbf{l}'} \langle \mathbf{l}\mathbf{K} | \hat{H}_0 | \mathbf{l}'\mathbf{K} \rangle F_{\lambda\mathbf{K}}(\mathbf{l}') = E_{\lambda\mathbf{K}} F_{\lambda\mathbf{K}}(\mathbf{l}), \quad (6.2)$$

где

$$\begin{aligned} \langle \mathbf{l}\mathbf{K} | \hat{H}_0 | \mathbf{l}'\mathbf{K} \rangle &= N^{-1} \sum_{\mathbf{k}\mathbf{k}'} e^{i\mathbf{k}\mathbf{l}' - i\mathbf{k}\mathbf{l}} \langle \mathbf{k}'\mathbf{K} | \hat{H}_0 | \mathbf{k}\mathbf{K} \rangle = \\ &= N^{-1} \sum_{\mathbf{k}} [E_c^{\text{HF}}(\mathbf{k}) - E_v^{\text{HF}}(\mathbf{k} - \mathbf{K})] e^{i\mathbf{k}(\mathbf{l} - \mathbf{l}')} - \\ &- \delta_{\mathbf{l}\mathbf{l}'} w_{\text{Coul}}(\mathbf{l}) + \delta_{\mathbf{l}\mathbf{0}} \delta_{\mathbf{l}'\mathbf{0}} \sum_{\mathbf{m}} e^{i\mathbf{k}\mathbf{m}} w_{\text{ex}}(\mathbf{m}). \end{aligned}$$

Здесь

$$w_{\text{Coul}}(\mathbf{l}) = \left\langle c\mathbf{l}, v\mathbf{0} \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| v\mathbf{0}, c\mathbf{l} \right\rangle$$

— энергия кулоновского взаимодействия между электроном и дыркой, максимумы волновых функций которых приходятся на ячейки, разделенные расстоянием \mathbf{l} (электрон имеет максимум волновой функции в ячейке \mathbf{l} , а дырка — в ячейке $\mathbf{0}$), а

$$w_{\text{exl}}(\mathbf{m}) = \left\langle c\mathbf{m}, v\mathbf{0} \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| c\mathbf{0}, v\mathbf{m} \right\rangle$$

— энергия обменного взаимодействия электрона и дырки для процесса, в котором в начальном состоянии как электрон, так и дырка находились в ячейке \mathbf{m} , а к конечному состоянию — в ячейке $\mathbf{0}$ (то есть возбужденное состояние переносится из ячейки \mathbf{m} в ячейку $\mathbf{0}$).

В общем виде уравнение (6.2) не может быть решено. Поэтому чаще всего ограничиваются двумя предельными случаями. Если электронно-дырочная пара связана слабо, что бывает в случае большой диэлектрической проницаемости и

сильной дисперсии $E_{\alpha}^{\text{HF}}(\mathbf{k})$, среднее расстояние между электроном и дыркой велико по сравнению с постоянной решетки, и такое состояние называется экситоном Ванье или экситоном большого радиуса [70]. В противоположном случае, когда электрон и дырка находятся в основном в одной и той же ячейке кристалла, имеют дело с экситоном Френкеля или с экситоном малого радиуса [72].

6.2 Экситон Френкеля

Если предположить, что радиус экситонного состояния мал, и электрон и дырка принадлежат в основном одной и той же ячейке кристалла, то можно положить $F_{\lambda\mathbf{K}}(\mathbf{l}) = \delta_{\mathbf{l}\mathbf{0}}$. При этом значение энергии получается непосредственно из секулярного уравнения

$$E_{\mathbf{K}} = \langle \mathbf{0K} | \hat{\mathcal{H}}_0 | \mathbf{0K} \rangle = E_c^0 - E_v^0 - w_{\text{Coul}}(\mathbf{0}) + w_{\text{ex}}(\mathbf{0}) + \sum_{\mathbf{m} \neq \mathbf{0}} w_{\text{ex}}(\mathbf{m}) e^{-i\mathbf{K}\mathbf{m}}, \quad (6.3)$$

где $E_{\alpha}^0 = N^{-1} \sum_{\mathbf{k}} E_{\alpha}^{\text{HF}}(\mathbf{k})$ – среднее по зоне Бриллюэна значение энергии одной из зон α (напомним, что такое рассмотрение справедливо для случая, когда энергетические зоны состоят только из одной ветви). Можно показать, что энергия $E_c^0 - E_v^0 - w_{\text{Coul}}(\mathbf{0}) + w_{\text{ex}}(\mathbf{0})$ соответствует энергии возбуждения изолированного атома (сравните с моделью сильно связанных электронов). В этом можно убедиться, увеличивая постоянную решетки. При этом сумма в последнем члене (6.3) стремится к нулю, и мы имеем дело с системой невзаимодействующих атомов. Последний член (6.3) отвечает резонансной передаче возбуждения из-за межатомного взаимодействия. Вся зависимость от полного момента $\hbar\mathbf{K}$ сосредоточена в этой сумме.

Для анализа этой зависимости предположим, что волновые функции сосредоточены вблизи отдельных атомов, и разложим в выражении для $w_{\text{ex}}(\mathbf{m})$ множитель $1/|\mathbf{r}_1 - \mathbf{r}_2|$ в ряд вблизи точек $\mathbf{r}_1 \approx \mathbf{m}$ и $\mathbf{r}_2 \approx \mathbf{m}$ (это можно сделать, поскольку член $\mathbf{m} = \mathbf{0}$ исключен из рассматриваемой суммы). Первый исчезающий член разложения приведет к выражению для диполь-дипольного взаимодействия дипольных моментов волновых функций, локализованных на атомах:

$$\omega_{\text{ex}}(\mathbf{m}) = \frac{|\mathbf{p}|^2 m^2 - 3(\mathbf{p}\mathbf{m})^2}{m^5},$$

где $\mathbf{p} = \langle \mathbf{c0} | \mathbf{er} | \mathbf{c0} \rangle$.

Обменная энергия ω_{ex} связана с вероятностью перехода возбуждения за счет диполь-дипольного взаимодействия из одной ячейки кристалла в другую, расположенную на расстоянии \mathbf{m} от первой. Характерное время резонансного перехода возбуждения составляет $\tau(\mathbf{m}) = 2\pi\hbar/\omega_{\text{ex}}(\mathbf{m})$. Следующие члены разложения дадут довольно быстро убывающие мультипольные взаимодействия. Суммы от $\omega_{\text{ex}}(\mathbf{m})$ могут быть аналитически вычислены в длинноволновом приближении $\mathbf{K} \rightarrow \mathbf{0}$ для ряда кристаллических решеток. Так, для кубических кристаллов имеем

$$\lim_{\mathbf{K} \rightarrow \mathbf{0}} \sum_{\mathbf{m}} \omega_{\text{ex}}(\mathbf{m}) e^{-i\mathbf{K}\mathbf{m}} = -\frac{4\pi}{3v} \left[p^2 - \frac{3(\mathbf{p}\mathbf{K})^2}{|\mathbf{K}|^2} \right], \quad (6.4)$$

где v — объем элементарной ячейки кристалла.

Как видно из (6.4), сдвиг энергии экситона Френкеля по отношению к энергии возбуждения изолированного атома имеет сингулярность при малых $\mathbf{K} \rightarrow \mathbf{0}$: его значение зависит от направления волнового вектора \mathbf{K} . Это связано с тем, что сумма (6.4) не имеет абсолютной сходимости ($\omega_{\text{ex}}(\mathbf{m})$ падает как m^{-3} , в то время как число членов с одинаковым m растет как площадь сферы с радиусом m , то есть как m^2). Если $\mathbf{K} \parallel \mathbf{p}$, экситон называется продольным, и его энергия в пределе $\mathbf{K} \rightarrow \mathbf{0}$ оказывается выше энергии атомного перехода: $E_{L0} = E_{\text{at}} + 8\pi p^2/3v$. В случае поперечного экситона, $\mathbf{K} \perp \mathbf{p}$, энергия ниже атомной: $E_{T0} = E_{\text{at}} - 4\pi p^2/3v$. Поскольку соотношение $\mathbf{K} \perp \mathbf{p}$ удовлетворяется для двух независимых направлений \mathbf{K} , существуют две ветви поперечных экситонов, которые в случае кубического кристалла оказываются вырожденными. С учетом этого вырождения «центр тяжести» всех трех ветвей совпадает с энергией перехода в атоме. Абсолютное значение разности энергий двух ветвей экситона при $\mathbf{K} = \mathbf{0}$ ($E_{L0} - E_{T0} = 4\pi p^2/v$) называется продольно-поперечным расщеплением, или LT -расщеплением [73]. Это расщепление имеет принципиальное значение и связано с появлением сингулярности в сумме по \mathbf{m} в выражении $\omega_{\text{ex}}(\mathbf{m})$, которая, в свою очередь, связана с дальнедействующим характером

кулоновского взаимодействия. В некубических кристаллах зависимость энергии экситона от направления вектора \mathbf{K} в пределе $\mathbf{K} \rightarrow \mathbf{0}$ также оказывается существенной.

Оптический переход под действием фотонов происходит как раз в состоянии с $\mathbf{K} = \mathbf{0}$, поскольку длина волны фотона, возбуждающего экситон, много меньше постоянной решетки. Однако переход фотона происходит лишь в состоянии поперечного экситона, поскольку электромагнитное поле поперечно, и возбуждаться могут переходы лишь в состояния, дипольный момент перехода в которые \mathbf{p} перпендикулярен волновому вектору. Это дает возможность определить положение линии оптического поглощения и соответствующую силу осциллятора. Используя общие формулы для $\epsilon(\omega)$, легко получить выражение для мнимой части тензора диэлектрической проницаемости $\epsilon_{ij}(\omega)$:

$$(\epsilon_2)_{ij}(\omega) = 4\pi^2 v^{-1} p_i p_j \delta(\hbar\omega - E_{T0}). \quad (6.5)$$

Однако форму линии поглощения и испускания света экситоном нельзя описать без привлечения представления о фонах. Это мы отложим для последующего изложения.

Переход в состояния продольных экситонов, а также возбуждение состояний с $\mathbf{K} \neq \mathbf{0}$, оказывается возможным лишь при взаимодействии с электромагнитным полем, возникающим в процессах рассеяния экситонов и зараженных частиц. В частности, в функцию потерь энергии $\text{Im}(-\tilde{\epsilon}^{-1})$ вносят вклад не поперечные, а только продольные экситоны. К этому вопросу можно будет вернуться при обсуждении экситонных поляритонов.

Выражение (6.5) не является изотропным. В кубических кристаллах необходимо учитывать три различных независимых дипольных момента \mathbf{p}^α (что связано с вырождением одной из участвующих в переходе зон; если валентная зона сформирована p -состояниями атомов, а зона проводимости — s -состояниями, такие переходы оказываются дипольно-разрешенными, и возникают три различных взаимно перпендикулярных дипольных момента, связанных с трехкратно вырожденными p -состояниями). Суммирование по всем α приводит к изотропному выражению для ϵ_2 :

$$\epsilon_2(\omega) = \frac{4\pi^2}{3v} p^2 \delta(\hbar\omega - E_{T0}). \quad (6.6)$$

Поведение действительной части диэлектрической проницаемости определяется соотношениями Крамерса — Кронига (1.21) (необходимо только принять во внимание, что формулы (6.5), (6.6) записаны только для положительных частот, и для отрицательных частот необходимо использовать соотношение (1.23)). Это приводит к появлению полюса в диэлектрической проницаемости:

$$\tilde{\varepsilon}(\omega) = \dots + \frac{8\pi E_{T0} p^2}{v[E_{T0}^2 - \hbar^2(\omega + i0)^2]}.$$

Несингулярный член в $\tilde{\varepsilon}(\omega)$, обозначенный в этой формуле точками, определяется из требования, что $\tilde{\varepsilon}(\omega)$ должно обращаться в ноль на частоте продольного экситона $\hbar\omega = E_{L0}$:

$$\tilde{\varepsilon}(\omega) = \varepsilon_{\infty}^{\text{ex}} + \frac{(\varepsilon_0^{\text{ex}} - \varepsilon_{\infty}^{\text{ex}})E_{T0}^2}{E_{T0}^2 - \hbar^2(\omega + i0)^2} \quad (6.7)$$

где введены выражения для диэлектрической проницаемости на частотах много больших частоты экситонного перехода

$$\varepsilon_{\infty}^{\text{ex}} = \frac{8\pi E_{T0} p^2}{v[E_{L0}^2 - E_{T0}^2]}$$

и на частотах много меньших частоты перехода (статическая диэлектрическая проницаемость):

$$\varepsilon_0^{\text{ex}} = \varepsilon_{\infty}^{\text{ex}} \frac{E_{L0}^2}{E_{T0}^2}.$$

Естественно, формула (6.7) справедлива только для изолированного экситонного перехода (только в этом случае вклад остальных переходов в соотношения Крамерса-Кронига может рассматриваться как постоянный), и только для таких переходов параметры $\varepsilon_0^{\text{ex}}$ и $\varepsilon_{\infty}^{\text{ex}}$ могут иметь определенный смысл.

Зависимость энергии экситона от величины волнового вектора \mathbf{K} приводит к образованию зоны экситонных состояний. Ширина этой зоны, так же как и LT -расщепление, определяется резонансным взаимодействием атомов в кристалле и по порядку величины равна $4\pi p^2/v$. При малых \mathbf{K} можно получить зависимость энергии экситона от \mathbf{K} в виде квадратичной формы от компонент импульса. Коэффициенты в этой

форме являются компонентами тензора обратной эффективной массы $(m^{-1})_{ij}$. Однако выражения для $(m^{-1})_{ij}$ имеют сложный вид и сильно зависят от типа кристаллической структуры. Знак эффективной массы может быть, как и в случае электронной зоны в приближении сильной связи, как положительным, так и отрицательным.

Приведенные выше замечания о числе ветвей экситона Френкеля справедливы в том случае, если в элементарной ячейке кристалла находится один атом (молекула), а у каждого атома имеется только одно возбужденное состояние. Если в элементарной ячейке кристалла имеются две или больше одинаковых молекул, то число ветвей увеличивается в соответствующее число раз, и между зонами возникает дополнительное расщепление. Величина такого расщепления определяется резонансным взаимодействием молекул в ячейке и по величине порядка LT -расщепления. Этот эффект называется «давидовским расщеплением» [74]. Этот эффект по своей природе аналогичен появлению оптической ветви фононных колебаний в кристаллах, содержащих два атома и более атомов в элементарной ячейке.

При описании экситонов Френкеля мы не пытались решать секулярное уравнение, а априори предположили, что волновая функция локализована только в одной ячейке. Применимость этой картины ограничена молекулярными кристаллами. Некоторые черты экситонов Френкеля имеются у нижних экситонных состояний в кристаллах инертных газов, щелочно-галогидных кристаллов и т.д. Но уже более возбужденные состояния являются более рыхлыми, и применение модели экситонов Френкеля уже неправильно. Эти состояния лучше описываются моделью экситонов Ванье.

6.3 Экситон Ванье-Мотта

Возбужденные состояния большого радиуса — другой предельный случай электронно-дырочных взаимодействий. Вернемся к секулярному уравнению (6.2).

Радиус экситона Френкеля порядка постоянной решетки. Тем самым, в соответствии с принципом неопределенности, в образовании такого состояния принимают участие состояния с импульсами из всей зоны Бриллюэна (что видно из сумм по \mathbf{k} , которые определяют энергетические характеристики экситонов Френкеля). Если предположить, что радиус экситонного состояния R много больше постоянной решетки a , то

согласно принципу неопределенности в формировании такого состояния должны принимать состояния с волновым вектором из области размером π/R , что много меньше, чем размеры зоны Бриллюэна π/a . Тем самым необходимо учитывать только состояния в небольшой окрестности экстремального волнового вектора $\mathbf{k}_m(\mathbf{K})$, для которого разность $E_c^{\text{HF}}(\mathbf{k}) - E_v^{\text{HF}}(\mathbf{k} - \mathbf{K})$ имеет минимум. Вблизи этого экстремума можно разложить эту разность (предполагается, что обе зоны невырождены вблизи экстремума):

$$E_c^{\text{HF}}(\mathbf{k}) - E_v^{\text{HF}}(\mathbf{k} - \mathbf{K}) = E_{\min}(\mathbf{K}) + \frac{\hbar^2}{2} \sum_{ik} (\mu^{-1}(\mathbf{K}))_{ik} (\mathbf{k} - \mathbf{k}_m)_i (\mathbf{k} - \mathbf{k}_m)_k + \dots,$$

где $(\mu^{-1}(\mathbf{K}))_{ik}$ — тензор приведенной обратной эффективной массы.

Кулоновский интеграл, входящий в секулярное уравнение, на больших расстояниях \mathbf{l} можно аппроксимировать обычным кулоновским взаимодействием двух зарядов с учетом влияния поляризации оставшейся части кристалла:

$$w_{\text{Coul}}(\mathbf{l}) = \left\langle c\mathbf{l}, v\mathbf{0} \left| \frac{e^2}{\epsilon|\mathbf{r}_1 - \mathbf{r}_2|} \right| v\mathbf{0}, c\mathbf{l} \right\rangle = \frac{e^2}{\epsilon R},$$

где R — непрерывная переменная, соответствующая дискретному вектору \mathbf{l} . Кроме того, поскольку вероятность нахождения электрона и дырки в одной точке мала, можно пренебречь обменным взаимодействием w_{ex} . Волновой вектор \mathbf{k} в континуальном приближении является оператором дифференцирования $\nabla = \partial/\partial\mathbf{R}$. В таком случае секулярное уравнение (6.2) может быть записано в следующем виде:

$$\left[-\frac{\hbar^2}{2} (\mu^{-1}(\mathbf{K}))_{ik} \nabla_i \nabla_k - \frac{e^2}{\epsilon R} \right] \tilde{F}_{\lambda\mathbf{K}}(\mathbf{R}) = \tilde{E}_{\lambda\mathbf{K}} \tilde{F}_{\lambda\mathbf{K}}(\mathbf{R}), \quad (6.8)$$

где $\tilde{E}_{\lambda\mathbf{K}} = E_{\lambda\mathbf{K}} - E_{\min}(\mathbf{K})$, $\tilde{F}_{\lambda\mathbf{K}}(\mathbf{R}) = F_{\lambda\mathbf{K}}(\mathbf{R}) e^{-i\mathbf{k}_m\mathbf{R}}$.

Приближение, в котором записано это уравнение, называется приближением эффективной массы. В случае невырожденных зон это уравнение определяет скалярную функцию $F_{\lambda\mathbf{K}}(\mathbf{R})$, которая является огибающей функций Ванье. В

случае вырожденных зон уравнение (6.8) заменяется на соответствующую систему уравнений.

Если рассматриваются малый волновой вектор \mathbf{K} , предполагается, что эффективные массы дырки $m_h = -m_v$ и электрона $m_c = m_e$ являются скалярными величинами, и что экстремумы обеих зон находятся в точке Γ ($\mathbf{k} = \mathbf{0}$), то

$$E_{\min}(\mathbf{K}) = \frac{\hbar^2 \mathbf{K}^2}{2(m_e + m_h)} + E_c^{\text{HF}}(\mathbf{0}) - E_v^{\text{HF}}(\mathbf{0}),$$

и приведенная эффективная масса $\mu^{-1} = m_e^{-1} + m_h^{-1}$ также является скалярной величиной и не зависит от \mathbf{K} . В этом простейшем случае уравнение (6.8) имеет такой же вид, что и уравнение атома водорода. Поэтому его собственные значения известны.

Если $\mu > 0$, то имеются связанные состояния с энергией $E_n = -Ry^*/n^2$, где $Ry^* = \mu e^4 / 2\varepsilon^2 \hbar^2$ — так называемая экситонная постоянная Ридберга. Эта константа отличается от обычной постоянной Ридберга для атома водорода $Ry = me^4 / 2\hbar^2 \approx 13,5$ эВ множителем μ/m_e^2 , где m — масса свободного электрона. Для ионизационного континуума (несвязанных состояний) энергия равна $E_{\mathbf{k}} = \hbar^2 k^2 / 2\mu$, где \mathbf{k} — импульс относительного движения электрона и дырки. Экситонный борковский радиус $a_B^* = \varepsilon \hbar^2 / \mu e^2$ в $\varepsilon m / \mu$ раз больше борковского радиуса атома водорода $a_B = \hbar^2 / me^2$ (0,05 нм). Радиус n -го состояния равен $R_n = n^2 a_B^*$.

Применимость модели экситона Ванье-Мотта требует, чтобы радиус состояния R_n был бы много больше постоянной решетки a . Это требует выполнения соотношения $n^2 \varepsilon m / \mu \gg 1$. Другими словами, модель Ванье-Мотта применима как для кристаллов с большим значением ε и малой приведенной эффективной массой, так и в случае больших значений квантового числа n (то есть для состояний вблизи порога ионизации).

Тем самым связанное состояние экситона Ванье-Мотта характеризуются полным импульсом экситона \mathbf{K} и квантовыми числами $\lambda = \{n, l, m\}$, соответствующими квантовым числам атома водорода. Состояния с $n > 1$ являются вырожденными как по орбитальному квантовому числу l ($0 \leq l \leq n-1$), так и по m ($-l \leq m \leq l$). Энергия связанного состояния зависит

только от \mathbf{K} и n :

$$E_{n\mathbf{K}} = E_g + \frac{\hbar^2 K^2}{2(m_e + m_h)} - \frac{\mu e^4}{\hbar^2 \epsilon^2 n^2}.$$

Здесь была введена ширина запрещенной зоны E_g , которая в использованных выше обозначениях выражается через энергии состояний Хартри-Фока:

$$E_g = E_c^{\text{HF}}(\mathbf{0}) - E_v^{\text{HF}}(\mathbf{0}).$$

Связанные состояния имеют точку сгущения у границы ионизации $E_{\min}(\mathbf{K})$. Бесконечное число уровней связанных состояний связано с природой потенциала взаимодействия между электроном и дыркой (неэкранированный кулоновский потенциал). Для более короткодействующих потенциалов число связанных состояний становится конечным и может быть даже равно нулю.

Ниже нам понадобится знать не только значения энергий, но и поведение волновой функции, в частности, поведение $F_{\lambda\mathbf{K}}$ вблизи нуля. Как следует из решения уравнения Шредингера для атома водорода, $F_{\lambda\mathbf{K}}(\mathbf{R}) \sim (R/a_B^*)^l$ для $R \rightarrow 0$. Отсюда можно заключить, что вероятность нахождения электрона и дырки в одной точке отлична от нуля только для s -состояний ($l = 0$).

Ионизованные состояния экситона образуют ионизационный континуум и могут характеризоваться либо полным импульсом экситона \mathbf{K} и импульсом относительного движения электрона и дырки \mathbf{k} , либо импульсами отдельных частиц \mathbf{k}_e и \mathbf{k}_h :

$$E_{\mathbf{k}\mathbf{K}} = E_g + \frac{\hbar^2 K^2}{2(m_e + m_h)} + \frac{\hbar^2 k^2}{2\mu} = E_g + \frac{\hbar^2 k_e^2}{2m_e} + \frac{\hbar^2 k_h^2}{2m_h}.$$

Энергия ионизованных состояний зависит от импульсов электрона и дырки так, как будто электрон и дырка с импульсами \mathbf{k}_e и \mathbf{k}_h не взаимодействуют друг с другом. Однако волновая функция $F_{\lambda\mathbf{K}}(\mathbf{R})$ отличается от волновой функции невзаимодействующих частиц. Это подробно будет рассмотрено ниже.

При рассмотрении экситонов Френкеля большое внимание уделялось LT -расщеплению. Для экситонов Ванье-Мотта также возможно LT -расщепление. Оно оказывается намного

слабее, поскольку определяется обменным взаимодействием электрона и дырки, которое возможно только для совпадающих координат электронов и дырок. Поэтому LT -расщепление оказывается порядка $4\pi p^2 |F_{n\mathbf{K}}(0)|^2 / v$ и мало, поскольку $F_{n\mathbf{K}}(0)$ мало и отличается от нуля только для s -состояний. Поэтому таким расщеплением обычно пренебрегают.

Рассмотрим оптические функции экситонов Ванье-Мотта. Впервые они были рассмотрены в [75]. Как и во всех случаях, когда рассматривается идеальная недеформируемая решетка, переход происходит в состояние полным волновым вектором $\mathbf{K} = 0$.

Матричный элемент перехода из основного состояния кристалла $|g\rangle$ в возбужденное состояние $|\lambda\mathbf{0}\rangle$ оказывается равным

$$\begin{aligned} e\langle\lambda\mathbf{0}|\mathbf{re}|g\rangle &= e\sum_{\mathbf{k}} f_{\lambda\mathbf{0}}^*(\mathbf{k}) \langle\mathbf{k}\mathbf{0}|\mathbf{re}|g\rangle = \\ &= e\sum_{\mathbf{k}} f_{\lambda\mathbf{0}}^*(\mathbf{k}) \langle g|a_{v\mathbf{k}}^+ a_{c\mathbf{k}}\mathbf{re}|g\rangle = e\sum_{\mathbf{k}} f_{\lambda\mathbf{0}}^*(\mathbf{k}) \int \psi_{c\mathbf{k}}^* \mathbf{re} \psi_{v\mathbf{k}} d^3r \equiv \\ &\equiv \sum_{\mathbf{k}} f_{\lambda\mathbf{0}}^*(\mathbf{k}) \mathbf{ed}_{cv}(\mathbf{k}). \end{aligned} \quad (6.9)$$

Воспользуемся тем, что $f_{\lambda\mathbf{0}}(\mathbf{k})$ для экситона большого радиуса отлична от нуля лишь в небольшой области зоны Бриллюэна вокруг значения $\mathbf{k} = \mathbf{k}_m(\mathbf{0})$, и разложим i -ю проекцию матричного элемента дипольного момента $\mathbf{d}_{cv}(\mathbf{k})$ в ряд по \mathbf{k} :

$$\begin{aligned} [\mathbf{d}_{cv}(\mathbf{k})]_i &= [\mathbf{d}_{cv}(\mathbf{k}_m(\mathbf{0}))]_i + \\ &+ [\mathbf{k} - \mathbf{k}_m(\mathbf{0})]_i [\nabla_{\mathbf{k}} [\mathbf{d}_{cv}(\mathbf{k})]_i]_{\mathbf{k}=\mathbf{k}_m(\mathbf{0})} + \dots \end{aligned}$$

Если $\mathbf{d}_{cv}(\mathbf{k}_m) \neq 0$, то переход называется дипольно разрешенным; в противном случае — $\mathbf{d}_{cv}(\mathbf{k}_m) = 0$ — дипольно запрещенным. В первом случае можно вынести $\mathbf{d}_{cv}(\mathbf{k}_m)$ за знак интеграла в (6.9), и для дипольного момента перехода получаем

$$e\langle\lambda\mathbf{0}|\mathbf{r}|g\rangle = \mathbf{d}_{cv}(\mathbf{k}_m) \sum_{\mathbf{k}} f_{\lambda\mathbf{0}}^*(\mathbf{k}) = \mathbf{d}_{cv}(\mathbf{k}_m) v^{-1/2} F_{\lambda\mathbf{0}}^*(\mathbf{0}).$$

Матричный элемент перехода отличен от нуля только в такие состояния, в которых $F_{\lambda\mathbf{0}} \neq 0$, т. е. только для s -состояний экситона Ванье-Мотта. Заметим, что это утверждение справедливо с точностью до параметра a/R_λ , где a —

постоянная решетки, R_λ — радиус экситона, поскольку, если этот параметр не является малым, выносить $\mathbf{d}_{cv}(\mathbf{k}_m)$ за знак суммы не представляется оправданным.

Для дипольно запрещенных переходов матричный элемент приобретает вид

$$\begin{aligned} e\langle \lambda \mathbf{0} | \mathbf{r} | g \rangle &= \sum_{\mathbf{k}} f_{\lambda \mathbf{0}}^*(\mathbf{k}) [[\mathbf{k} - \mathbf{k}_m] \nabla_{\mathbf{k}}] \mathbf{d}_{cv}(\mathbf{k}) \Big|_{\mathbf{k}=\mathbf{k}_m} = \\ &= v^{-1/2} \sum_i \frac{\partial \mathbf{d}}{\partial k_i} \Big|_{\mathbf{k}=\mathbf{k}_m} \frac{\partial F_{\lambda \mathbf{0}}^*}{\partial R_i} \Big|_{\mathbf{R}=0}. \end{aligned}$$

В этом случае переходы происходят не в состояния s -типа, а p -типа, поскольку только для таких состояний $\nabla F_{\lambda \mathbf{0}}$ отлично от нуля в точке $\mathbf{R} = 0$. Заметим, что термин «дипольно запрещенные переходы» имеет в спектроскопии твердого тела смысл, отличный от смысла в атомной и молекулярной спектроскопии. В последнем случае, если переход дипольно запрещен, возможны только магнитно-дипольные и квадрупольные переходы. В нашем же случае продолжает работать дипольное приближение, но для других областей зоны Бриллюэна. Если радиус экситона не слишком велик по сравнению с постоянной решетки, интенсивность дипольно запрещенных переходов не пренебрежимо мала по сравнению с дипольно разрешенными переходами.

Мнимая часть диэлектрической проницаемости для случая дипольно разрешенных переходов равна

$$\varepsilon_2(\omega) = \frac{4\pi^2}{v} |\mathbf{d}_{cv}|^2 \sum_{\lambda} |F_{\lambda \mathbf{0}}(\mathbf{0})|^2 \delta(\hbar\omega - E_{\lambda \mathbf{0}}).$$

Для связанных состояний при энергиях $\hbar\omega = E_{n\mathbf{0}} = E_g - Ry^*/n^2$ появляются дискретные линии с интенсивностями, пропорциональными $|F_{n\mathbf{0}}(\mathbf{0})|^2 = v/\pi a_B^{*3} n^3$, где a_B^* — боровский радиус экситона. Это означает, что интенсивности переходов в $1s$ -, $2s$ - и $3s$ -состояния экситона относятся как $1:1/8:1/27$, то есть их интенсивности быстро спадают. В приближении недеформируемой решетки все линии являются бесконечно узкими; в действительности они приобретают конечную ширину за счет взаимодействия с другими квазичастицами, в частности, фононами. Поэтому линии

высоковозбужденных связанных состояний ($n > 3$) размываются и образуют квазинепрерывный спектр с конечной интенсивностью. Хотя интенсивность каждой отдельной линии падает как n^{-3} , интенсивность квазинепрерывного спектра оказывается конечной, поскольку число состояний, попадающих интервал энергии ΔE , равно $(dE/dn)^{-1} \Delta E = n^3 \Delta E / 2Ry^*$. Отсюда следует, что интенсивность экситонного поглощения для больших n не зависит от энергии и равна $v/2\pi Ry^* a_B^*$ вплоть до E_g , то есть энергии ионизации экситона.

Выше порога ионизации энергия экситонного состояния совпадает с энергией невзаимодействующей электронно-дырочной пары. Но взаимодействие электрона и дырки существенно меняет волновую функцию $F_{\mathbf{k}\mathbf{0}}(\mathbf{0})$. Для водородоподобных состояний с энергией выше E_g вероятность нахождения электрона и дырки в одной точке $|F_{\mathbf{k}\mathbf{0}}(\mathbf{0})|^2$ равна [49] $v\pi\alpha \exp(\pi\alpha)/\text{sh}(\pi\alpha)$, где $\alpha = (a_B^*k)^{-1}$ и $k = \hbar^{-1}[2\mu(\hbar\omega - E_g)]^{1/2}$; при этом $\alpha \rightarrow \infty$ для $\hbar\omega \rightarrow E_g$, то есть эта вероятность растет с приближением энергии к порогу ионизации сверху как $(\hbar\omega - E_g)^{-1/2}$, что полностью компенсирует уменьшение приведенной межзонной плотности состояний по закону $(\hbar\omega - E_g)^{1/2}$. В пределе $\hbar\omega \rightarrow E_g$ со стороны континуума значение ε_2 равно

$$\begin{aligned} \varepsilon_2 &\approx \frac{1}{(2\pi)^3} \int d^3k \frac{\pi v}{a_B^* k} \delta\left(\hbar\omega - E_g - \frac{\hbar^2 k^2}{2\mu}\right) = \\ &= \frac{v}{2\pi a_B^*} \int_0^\infty k dk \frac{\delta\left(k - \sqrt{2\mu(\hbar\omega - E_g)/\hbar}\right)}{\hbar^2 k/\mu} = \\ &= \frac{v\mu}{2\pi a_B^* \hbar^2} = \frac{v}{2\pi Ry^* a_B^{*3}}. \end{aligned}$$

Таким образом, спектры $\varepsilon_2(\omega)$ выше и ниже E_g плавно, безо всяких особенностей переходят друг в друга (рис. 20а). Тем самым при наличии сильных экситонных эффектов особенности Ван Хова $\varepsilon_2(\omega) \sim (\hbar\omega - E_g)^{1/2}$ не наблюдаются совсем, то есть граница между связанными и ионизованными состояниями никак не проявляется в спектрах $\tilde{\varepsilon}(\omega)$ и в спектрах

поглощения. Влияние экситонных эффектов ощущается в области $2Ry^*$ выше E_g и только при $\hbar\omega - E_g \gg Ry^*$ $\varepsilon_2(\omega)$ становится пропорциональным $(\hbar\omega - E_g)^{1/2}$.

В случае дипольно запрещенных переходов нижним состоянием, возбуждаемым светом, является $2p$ -состояние экситона, и интенсивности переходов пропорциональны

$$\frac{v}{\pi a_B^*5} \frac{n^2 - 1}{n^5}$$

в области дискретных переходов, и пропорциональны

$$|\nabla_{\mathbf{R}} F_{\mathbf{k}\mathbf{0}}|_{\mathbf{R}=\mathbf{0}}|^2 = v\pi\alpha^2(1 + \alpha^2) \exp(\pi\alpha)/\text{sh}(\pi\alpha)$$

в области непрерывного спектра. Как и в предыдущем случае, квазинепрерывный спектр для $\hbar\omega < E_g$ без сингулярностей переходит в непрерывный для $\hbar\omega > E_g$:

$$\varepsilon_2(\omega) = 2\pi \left| \frac{\partial \mathbf{d}_{cv}}{\partial \mathbf{k}} \right|^2 \frac{\hbar\omega - E_g + Ry^*}{Ry^{*2} a_B^*5}, \text{ при } |\hbar\omega - E_g| \ll Ry^*.$$

Особенности Ван Хова не наблюдаются и в этом случае. Качественно форма спектра с учетом уширения линий показана на рис. 20б. Интенсивности первых двух линий, соответствующих переходам в $2p$ - и $3p$ -состояния, относятся как $(3/32):(8/243) = 3^6/2^8 \approx 3:1$.

Для многих полупроводников экситонная постоянная Ридберга мала, например в германии $Ry^* \approx 0,016$ эВ [76]. Кроме того, экситонный радиус в Ge $a_B^* \sim 4$ нм и значительно превышает постоянную решетки. Поэтому в полупроводниках, с одной стороны, хорошо выполняются условия применимости модели экситонов Ванье–Мотта, а, с другой

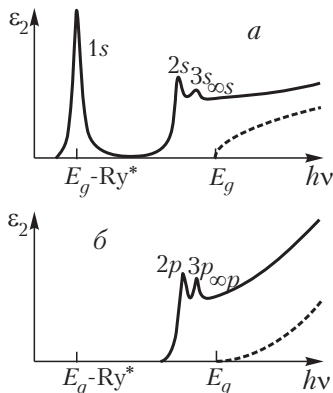


Рис. 20. Экситонный спектр поглощения по модели Эллиота-Ванье в случае разрешенных (а) и запрещенных (б) переходов. Пунктиром изображены кривые ε_2 без учета экситонных эффектов.

стороны, они слабо проявляются в оптическом поглощении. Напротив, в более широкозонных кристаллах Ry^* возрастает. Так, в Cu_2O — веществе, в котором впервые наблюдалось большое число экситонных линий, соответствующих большим n , Ry^* порядка 0,1 эВ [77]; а в щелочно-галогидных кристаллах достигает 0,5 эВ. Соответственно уменьшается и радиус состояния. Поэтому в этих кристаллах область влияния экситонных состояний составляет 1,5–2 эВ, и особенности Ван Хофа в принципе наблюдать невозможно.

Здесь необходимо сделать следующее замечание. Для оценки ширины запрещенной зоны E_g во многих кристаллах, в том числе и в широкозонных, часто используют модель экситонов Ванье—Мотта, причем ограничиваются иногда первыми двумя-тремя пиками экситонной серии, получая из них значение Ry^* . При этом не учитывается, что $1s$ -состояние имеет малый радиус и к нему модель экситона большого радиуса обычно плохо применима. Поэтому к такого рода оценкам ширины запрещенной зоны следует относиться с осторожностью. Определенным критерием применимости модели Ванье—Мотта может служить отношение площадей под $1s$ - и $2s$ -пиками, которое должно составлять 8:1.

Здесь рассмотрены основные черты экситонных эффектов, которые проявляются вблизи минимума $E_c^{HF} - E_v^{HF}$ — особой точки M_0 -типа по классификации §5. В случае неизотропных, вырожденных зон расчет экситонных эффектов значительно усложняется, однако основные характеристики ε_2 меняются мало. Вблизи остальных критических точек (M_1 , M_2 , M_3) тоже возможно возникновение экситонных эффектов, которые приводят к изменению формы спектра поглощения по сравнению с рассмотренным в предыдущем параграфе. Так, в критической точке M_3 электрон и дырка эффективно отталкиваются друг от друга, что приводит к уменьшению вероятности их нахождения в одной точке и, следовательно, к уменьшению вклада от соответствующей пары зон в ε_2 . Естественно, что в этой точке не образуется связанных состояний. Величина $|F_{\lambda_0}(\mathbf{0})|^2$ уменьшается экспоненциально быстро (как $e^{-1/\sqrt{x}}$) при приближении к точке M_3 ($x \sim E_{M_3} - \hbar\omega$), и поэтому вклад в $\varepsilon_2(\omega)$ плавно переходит в ноль.

Увеличение $\varepsilon_2(\omega)$ при энергиях $\hbar\omega$ вокруг точки M_0 и их уменьшение ниже точки M_3 связаны друг с другом,

поскольку $\varepsilon_2(\omega)$ должны удовлетворять правилу сумм. Тем самым, усиление экситонных эффектов ниже порога E_g с необходимостью приводит к перераспределению сил осцилляторов во всем спектре. В пределе очень сильной электрон-фононной связи все силы осцилляторов концентрируются в одной линии, которая в этом пределе описывается моделью экситона Френкеля. Силы осцилляторов для переходов в состоянии непрерывного спектра становятся малыми.

В критических точках M_1 и M_2 в зависимости от соотношения эффективных масс также возможно существование резонансных состояний. Эти состояния приводят к пикам в поглощении. Однако, поскольку энергия этих состояний лежит выше минимальной энергии возбуждения основного состояния, состояния будут автоионизационными — они распадаются на другие электронные возбуждения. Это приводит к конечной ширине соответствующих энергетических уровней даже в недеформируемом кристалле. Методы расчета форм линий подобных состояний изложены в следующем параграфе.

6.4 Экситоны в реальных кристаллах

Выше были рассмотрены две крайние модели: модели экситона Френкеля и экситона Ванье-Мотта. Естественно, что в большинстве случаев для реальных кристаллов эти простые модели в чистом виде неприменимы. В данном разделе мы рассмотрим некоторые эффекты, опущенные в предыдущих разделах.

Сделаем несколько замечаний по поводу возможных спиновых состояний экситона ([71, 78]). Поскольку электрон и дырка имеют спин, равный $\frac{1}{2}$, то возможно возникновение возбужденного состояния с полным спином $S = 0$ или $S = 1$. В первом случае экситон называется синглетным экситоном, во втором — триплетным. При оптическом возбуждении спин электрона при переходе из валентной зоны в зону проводимости не меняется. Исходное состояние $|g\rangle$ обладает равным нулю спином, поэтому при взаимодействии с фотонами наблюдаются только синглетные экситоны. В то же время триплетные экситоны играют значительную роль в процессах переноса энергии и в люминесценции (особенно в молекулярных кристаллах). При выводе выражений для энергии экситонов необходимо к координатам приписать спиновые компоненты и везде, где входит интегрирование по \mathbf{r} , производить

суммирование по спину соответствующего электрона. При этом от спинового состояния экситона будет зависеть обменный член. Это приводит к тому, что для триплетных состояний w_{ex} в уравнении (6.3) исчезает, а для синглетных — удваивается. Поскольку $w_{\text{ex}}(0) > 0$, синглетное состояние лежит выше по энергии, чем триплетное, на величину $2 \sum_{\mathbf{m} \neq 0} e^{i\mathbf{K}\mathbf{m}} w_{\text{ex}}(\mathbf{m})$. Обменное взаимодействие является кон-
тактным, поэтому оно существенно только в s -состояниях экситона и пропорционально $(a/a_B^*)^3$.

Здесь был рассмотрен спектр экситонных состояний в случае, когда зона проводимости и валентная зона состоят из одной ветви каждая. Однако часто валентная зона в точке максимума — в Γ -точке — является вырожденной. Спин-орбитальное взаимодействие частично снимает это вырождение, и валентная зона расщепляется на подзоны. Это приводит к тому, что во многих кристаллах наблюдаются две-три экситонные серии вблизи края фундаментального поглощения, причем они имеют различные постоянные Ридберга и различные энергии точек сгущения квазидискретного спектра. Наблюдение нескольких серий позволяет с большой достоверностью говорить о структуре валентной зоны вблизи максимума.

Использование метода эффективной массы и аппроксимация кулоновского взаимодействия введением макроскопической диэлектрической проницаемости возможны лишь для потенциалов, плавно меняющихся на расстояниях порядка межатомных. Для кулоновского потенциала вблизи $r = 0$ это условие не выполняется, поэтому даже в кристаллах с хорошо выраженными водородоподобными сериями первый уровень $1s$ выпадает из общей закономерности. Чтобы понять, как можно откорректировать потенциал, требуется подробнее обсудить поляризацию окружающего кристалла.

Продольное электрическое поле, возникающее при взаимодействии электрона с дыркой, экранируется в кристалле, что учитывается диэлектрической проницаемостью ϵ в знаменателе кулоновского взаимодействия. Поскольку ϵ зависит от частоты электрического поля ω , поляризация кристалла также будет зависеть от характерных частот относительного движения электрона и дырки. Если радиус экситонного состояния мал (порядка межатомного расстояния), характерные частоты много больше частот ионной подсистемы (мы будем

использовать обозначение Ω_L для характерной частоты фононов), и в качестве диэлектрической проницаемости выступает ε_∞ , соответствующая области прозрачности диэлектрика. Для больших радиусов экситона характерные частоты движения становятся меньше, и в поляризации начинает участвовать ионная подсистема кристалла. При этом в качестве диэлектрической проницаемости выступает статическая диэлектрическая проницаемость ε_0 , соответствующая области частот ниже частоты инфракрасного поглощения ($\omega < \Omega_L$). Вообще говоря, из интуитивных соображений можно записать потенциал взаимодействия в виде

$$\omega_{\text{Coul}} = \frac{e^2}{\varepsilon(\omega_{eh}(r))r}, \quad (6.10)$$

где $\omega_{eh}(r)$ — характерная частота движения электрона и дырки, находящихся на расстоянии r друг от друга. Ее можно оценить, зная, что угловой момент порядка \hbar , а момент инерции порядка μr^2 : $\omega_{eh}(r) \approx \hbar/\mu r^2$. Для кристаллов с широкой запрещенной зоной эта частота становится сравнимой с типичной частотой колебаний решетки Ω_L на расстояниях r порядка нескольких нанометров. В работе [79] предложен более прямой способ построения интерполяционной формулы, описывающей переход от потенциала $e^2/\varepsilon_0 r$ на больших расстояниях к потенциалу $e^2/\varepsilon_\infty r$ для малых расстояний. Эта формула имеет вид (6.10), где $\varepsilon(r)$ определяется уравнением

$$\frac{1}{\varepsilon(r)} = \frac{1}{\varepsilon_\infty} - \left[\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right] \left(1 - \frac{e^{-r/\rho_e} + e^{-r/\rho_h}}{2} \right), \quad (6.11)$$

а $\rho_\alpha = \sqrt{\hbar/2m_\alpha\Omega_L}$. Интерполяционные формулы такого типа могут применяться и для других форм описания частотной зависимости $\varepsilon(\omega)$. Например, в [80] рассмотрено влияние локализованных экситонных состояний (френкелевских экситонов), имеющих энергию возбуждения E_{L0} , на поляризацию кристалла, и получена формула (6.11) с заменой $\varepsilon_0 \rightarrow \varepsilon_\infty$, $\varepsilon_\infty \rightarrow 1$, $\Omega_L \rightarrow E_{L0}/\hbar$. Взаимодействие электрона и дырки с френкелевскими экситонами имеет место только для больших характерных частот движения, то есть для расстояний, сравнимых с межатомными. На таких расстояниях применимость

континуального приближения уже вызывает сомнения.

Изменение формы кулоновского потенциала взаимодействия w_{Coul} приводит к невозможности получения аналитического решения уравнения для волновой функции экситона. Исследование экситонных состояний с потенциалом типа (6.11) возможно только численно. Но качественно свойства экситонов не меняются. В частности, остаются такие свойства, как бесконечное количество связанных состояний и пропадание особенностей Ван Хофа в спектре поглощения.

Если в кристалле имеются свободные носители заряда (и поэтому у него имеется конечная проводимость σ), возникает эффект экранирования кулоновского потенциала, природа которого также связана с поляризацией кристалла (учитывая, что на малых частотах диэлектрическая проницаемость имеет полюс: $\tilde{\epsilon}(\omega) = \epsilon_0 + 4\pi i\sigma/\omega$, см. (1.22)). Если мы предположим, что носители заряда не испытывают соударений, σ в свою очередь становится мнимой величиной $\sigma = ie^2 n/m^* \omega$ (где n — концентрация свободных зарядов, m^* — их эффективная масса), и диэлектрическая проницаемость приобретает вид $\epsilon(\omega) = \epsilon_0 - 4\pi e^2 n/m^* \omega^2$ и тем самым имеет полюс второго порядка при $\omega \rightarrow 0$ (как это было бы в случае сверхпроводника; сравните с (4.27)). Более прямой вывод выражения для диэлектрической проницаемости показывает, что нельзя пренебрегать пространственной дисперсией при наличии свободных зарядов, и $\epsilon(\omega, \mathbf{k})$ имеет также полюс и по \mathbf{k} в точке $\omega = 0$: $\epsilon(0, \mathbf{k}) = \epsilon_0 + 1/\rho_D^2 k^2$, где введен дебаевский радиус $\rho_D = \sqrt{k_B T / 4\pi n e^2}$ для невырожденного электронного газа. Такая зависимость диэлектрической проницаемости от \mathbf{k} приводит к экранированию продольного электромагнитного взаимодействия:

$$w_{\text{Coul}} = \frac{e^2}{r} e^{-r/\rho_D}. \quad (6.12)$$

В металлах такое экранирование является особенно сильным, и радиус экранирования ρ_{TF} , называемый радиусом Томаса-Ферми, определяется квантовой формулой и не зависит от температуры: $\rho_{\text{TF}} = \sqrt{\pi \hbar^2 / 4m^* e^2 k_F}$, где k_F — импульс Ферми. Такое экранирование приводит не только к сдвигу экситонных состояний, но и к изменению их числа. Поскольку

потенциал (6.12) становится короткодействующим, число связанных состояний конечно. В случае больших σ (например, в металлах) вовсе не образуется экситонных состояний. В случае конечного числа уровней сингулярности Ван Хофа должны наблюдаться на пороге фундаментального поглощения.

В ионных кристаллах и ряде других соединений экситоны имеют достаточно малый радиус. Но в их формировании принимают участие волновые функции не одного атома, а нескольких, поэтому модель экситона Френкеля также не применима. Для описания таких экситонов была предложена модель «переноса электрона», при которой электрон с атома галоида (валентная зона в основном сформирована p -состояниями аниона) переходит на атом щелочного металла (дно зоны проводимости формируется s - и d -функциями катиона). Энергия такого переноса по модели Хилша — Поля [81] равна

$$E = E_A - E_I + \alpha_M e^2 / a, \quad (6.13)$$

где E_A — энергия сродства электрона к атому галоида (энергия, требуемая для удаления электрона), E_I — энергия ионизации атома щелочного металла (которая выделяется при возвращении электрона на соседний атом), α_M — константа Маделунга, a — расстояние между ближайшими соседями. Последний член учитывает ионную структуру кристалла. Константа Маделунга определяется следующим образом. Потенциал в точке расположения какого-либо положительного иона в решетке из периодически чередующихся ионов с зарядами $\pm e$ из соображений размерности может зависеть только от типа решетки и постоянной решетки. В ионных кристаллах кубической структуры вместо постоянной решетки в качестве параметра часто используют расстояние между двумя ближайшими соседями a . Таким образом, этот потенциал имеет вид $\alpha_M e^2 / a$, где α_M зависит только от типа кристалла. Для решетки типа NaCl $\alpha_M = 1,7476$, для CsI $\alpha_M = 1,7624$.

Формула (6.13), дает достаточно точное положение первого пика поглощения. Эта формула правильно дает расщепление экситонных пиков из-за спин-орбитального взаимодействия, поскольку энергия сродства электрона к атому галоида принимает два значения в зависимости от состояния, в котором находится p^5 -оболочка атома: ${}^2P_{1/2}$ или ${}^2P_{3/2}$. К

сожалению, формула Хилша—Поля не может быть просто обоснована. Все попытки вывести ее с большей степенью строгости приводили к худшему согласию с экспериментом, в частности из-за того, что волновая функция электрона довольно размазана и вероятность найти электрон вне центрального атома и его шести ближайших соседей оказывается существенной. К тому же формулы этого типа дают только энергию нижнего возбужденного состояния, оценить же энергию ионизации такого экситона не удастся.

Экситонные состояния возникают не только при переходе из валентной зоны в зону проводимости, но и при переходе из более глубоко лежащих по энергии остовных зон. При этом основные свойства экситона аналогичны валентным экситонам, за исключением того, что из-за слабого перекрытия внутренних электронных функций дисперсия остовных зон мала. При этом полная масса экситона оказывается большой, а приведенная масса относительного движения электрона и дырки определяется только зоной проводимости. Однако в этих состояниях возможно взаимодействие с менее энергетичными состояниями, что приводит к конечному времени жизни экситонов. Эти эффекты рассматриваются в следующем параграфе.

§7 Конфигурационное взаимодействие

7.1 Диэлектрическая проницаемость системы взаимодействующих частиц

В параграфах 5 и 6 были получены выражения для диэлектрической проницаемости $\tilde{\epsilon}(\omega)$ в предположении, что известны точные решения $|\alpha\rangle$ уравнения Шредингера и соответствующие им собственные значения E_α . К сожалению, в большинстве случаев известны только приближенные собственные функции и собственные решения гамильтониана $\hat{\mathcal{H}}_0$, описывающего твердое тело без взаимодействия с фотонами. Если же известны только собственные функции и собственные значения некоторого приближенного гамильтониана $\hat{\mathcal{H}}_{00}$, то такая система функций может быть также использована для решения точной задачи, поскольку используемая система функций является полной. Однако гамильтониан $\hat{\mathcal{H}}_0$ не диагонализуются в этой системе функций. Тем самым состояние, описываемое собственной функцией гамильтониана $\hat{\mathcal{H}}_{00}$, приобретает конечное время жизни в системе, описываемой точным гамильтонианом $\hat{\mathcal{H}}_0$, а соответствующая энергия приобретает некоторую мнимую часть. При этом формула (3.8) для $\epsilon_2(\omega)$ теряет смысл. Попытаемся обобщить эту формулу не очень строго, но наглядным способом.

Мнимая часть диэлектрической проницаемости пропорциональна сумме по конечным состояниям:

$$\begin{aligned}\epsilon_2(\omega) &\sim \sum_f \left| \langle f | \hat{\mathcal{M}} | g \rangle \right|^2 \delta(\hbar\omega - \hbar\omega_{fg}) = \\ &= \sum_f \langle g | \hat{\mathcal{M}}^+ | f \rangle \delta(\hbar\omega - \hbar\omega_{fg}) \langle f | \hat{\mathcal{M}} | g \rangle.\end{aligned}\tag{7.1}$$

Мы предположили, что начальное состояние является

основным состоянием $|g\rangle$ нашей системы и не вырождено. Введенный здесь для краткости оператор электрон-фотонного взаимодействия $\hat{M} \sim \sum_j \mathbf{r}_j \mathbf{e}$ является эрмитовым, $\hat{M}^+ = \hat{M}$.

Заменим входящую в (7.1) δ -функцию эквивалентным выражением

$$\delta(\hbar\omega - \hbar\omega_{fg}) = \frac{1}{\pi} \text{Im} \frac{1}{\hbar\omega - E_f + E_0 - i0},$$

где E_0 — энергия основного состояния. Если энергию отсчитывать от энергии основного состояния, то $E_0 = 0$ и можно переписать формулу (7.1) следующим образом:

$$\varepsilon_2(\omega) \sim \pi^{-1} \text{Im} \sum_f \langle g | \hat{M} | f \rangle \langle f | \hat{G} \hat{M} | g \rangle,$$

$$\hat{G} = (\hbar\omega - \hat{H}_0 - i0)^{-1}. \quad (7.2)$$

Здесь мы ввели функцию Грина \hat{G} для точного гамильтониана \hat{H}_0 и учли, что в состоянии $|f\rangle$ его значение равно E_f . Поскольку система функций $|f\rangle$ является полной, можно воспользоваться равенством $\sum_f |f\rangle \langle f| = 1$ и окончательно получить

$$\varepsilon_2(\omega) \sim \pi^{-1} \text{Im} \langle g | \hat{M} \hat{G} \hat{M} | g \rangle. \quad (7.3)$$

Эта формула не включает в себя конечные состояния явным образом. Поэтому такое выражение может быть вычислено для любой полной системы функций, например, для системы собственных функций невозмущенного гамильтониана \hat{H}_{00} . Формула (7.3) может быть выведена вполне строго. Для того чтобы ею можно было воспользоваться, выясним некоторые свойства функции Грина. Пусть гамильтониан \hat{H}_0 представлен в виде «невозмущенного» гамильтониана \hat{H}_{00} и некоторого возмущения $\hat{V} = \hat{H}_0 - \hat{H}_{00}$, которое включает в себя взаимодействия между различными состояниями гамильтониана \hat{H}_{00} . Для функции Грина может быть записано так называемое «уравнение Дайсона»:

$$\hat{G} = \hat{G}_0 + \hat{G} \hat{V} \hat{G}_0, \quad (7.4)$$

где $\hat{\mathcal{G}}_0 = (\hbar\omega - \hat{\mathcal{H}}_{00} - i0)^{-1}$ — функция Грина невозмущенной задачи. Уравнение Дайсона может быть легко проверено умножением справа на $\hat{\mathcal{G}}_0^{-1}$ и слева на $\hat{\mathcal{G}}^{-1}$, после чего получается тождество

$$\hbar\omega - i0 - \hat{\mathcal{H}}_{00} \equiv \hat{\mathcal{G}}_0^{-1} = \hat{\mathcal{G}}^{-1} + \hat{\mathcal{V}}.$$

Подчеркнем, что уравнение (7.4) точное. К сожалению, поскольку уравнение операторное, то решить его в общем виде невозможно (за исключением формального решения (7.2), мало дающего для практики). Большинство методов решения уравнения Дайсона основано на переписывании его в виде бесконечного ряда:

$$\hat{\mathcal{G}} = \hat{\mathcal{G}}_0 + \hat{\mathcal{G}}_0 \hat{\mathcal{V}} \hat{\mathcal{G}}_0 + \hat{\mathcal{G}}_0 \hat{\mathcal{V}} \hat{\mathcal{G}}_0 \hat{\mathcal{V}} \hat{\mathcal{G}}_0 + \hat{\mathcal{G}}_0 \hat{\mathcal{V}} \hat{\mathcal{G}}_0 \hat{\mathcal{V}} \hat{\mathcal{G}}_0 \hat{\mathcal{V}} \hat{\mathcal{G}}_0 + \dots \quad (7.5)$$

Этот ряд для функции Грина используется для приближенно исследования систем взаимодействующих частиц.

7.2 Эффект Оже

В многоэлектронной системе при достаточно высоких энергиях возбужденных состояний возникает особый тип вырождения, когда одно и то же значение энергии соответствует двум и более состояниям, имеющим различное число возбужденных частиц. Например, если электрон извлекается из одной из достаточно глубоких остовных зон, энергия такого состояния может быть равна энергии более сложного состояния, в котором два электрона перемещаются из валентной зоны на возбужденные уровни. Такие состояния с разным числом возбужденных электронов могут взаимодействовать друг с другом, что вызывает переходы между ними. Поскольку состояний такого рода может быть много (электрон может перейти в некоторое состояние в зоне проводимости, образовать экситон, перейти на дефектное состояние и т. д.), подробный анализ всех возможных ситуаций достаточно громоздок. Здесь мы опишем одну из таких возможностей, но основные черты такого описания будут применимы и для других случаев. Рассматриваемые процессы взаимодействия называются эффектом Оже (см., например, [67, 82, 83]).

Рассмотрим систему уровней в твердом теле, показанную на рис. 21. Предположим, что имеются некоторые заполненные в основном состоянии остовные уровни c_1 , c_2 и c_3 . Выше

энергии E_t начинается континуум непрерывных состояний. Электромагнитное излучение переводит электрон с самого глубокого основного уровня c_1 в некоторое возбужденное состояние. Таким состоянием может быть состояние в зоне проводимости или экситонное состояние α . Здесь будет рассмотрен последний случай, поскольку при отсутствии взаимодействия такой переход приводит к появлению изолированной δ -образной линии в спектре поглощения.

Будем считать, что под ионизационным континуумом имеется свободное состояние α , в которое электрон переходит с уровня c_1 . Для обозначения такого состояния будем использовать операторы рождения и уничтожения: $|\alpha\rangle = a_\alpha^+ a_{c_1} |g\rangle$. На рис. 21 этот процесс показан штриховой линией. Если же учитывать взаимодействие с другими состояниями, то выписанное выше состояние уже не является собственным состоянием точного гамильтониана \hat{H}_0 , и его время жизни становится конечным. Будем предполагать, что приближенному гамильтониану \hat{H}_{00} соответствуют невзаимодействующие состояния c_1, c_2, c_3, α и состояния в зоне проводимости k . Взаимодействие \hat{V} приводит к частичному перемешиванию этих состояний. В нашей системе энергию E_α , которую имеет состояние $|\alpha\rangle$, может иметь еще несколько состояний. Например, таким состоянием может быть состояние с двумя возбужденными электронами $|\alpha\rangle = a_k^+ a_\alpha^+ a_{c_2} a_{c_3} |g\rangle$, в котором дырки возникают на менее глубоких уровнях. Это возможно в том случае, если энергия порога ионизационных состояний удовлетворяет соотношению $E_t \leq E_{c_2} + E_{c_3} - E_{c_1}$. При этом возникновение состояния $|k\rangle$ из состояния $|\alpha\rangle$ можно описать следующим образом: электрон с одного из уровней, например c_2 , заполняет дырку в состоянии c_1 . При этом излучается продольный фотон – квант кулоновского поля. Этот фотон сразу же поглощается электроном, находящимся в состоянии c_3 , возбуждая его на один из уровней континуума k . Именно этот процесс и называется процессом Оже. В той же схеме уровней в состоянии

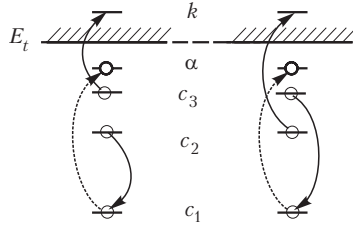


Рис 21. Возможные схемы Оже-переходов в твердом теле. Радиационные переходы показаны пунктирными стрелками.

c_1 может перейти электрон из состояния c_3 , а в состояние континуума – электрон из состояния c_2 . Такая схема процессов показана на правой половине рис. 21. Оба эти процесса приводят к одинаковому конечному состоянию.

Матричный элемент взаимодействия между двумя состояниями $|\alpha\rangle$ и $|k\rangle$ можно записать в виде

$$\begin{aligned} \mathcal{V}_{k\alpha} &= \left\langle \alpha \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| k \right\rangle = \left\langle g \left| a_{c_3}^+ a_{c_2}^+ a_{\alpha} a_k \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} a_{\alpha}^+ a_{c_1} \right| g \right\rangle = \\ &= \int d^3 r_1 d^3 r_2 \psi_{c_1}^*(\mathbf{r}_1) \psi_{c_2}^*(\mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \times \\ &\times \left[\psi_{c_2}(\mathbf{r}_2) \psi_{c_3}(\mathbf{r}_1) - \psi_{c_3}(\mathbf{r}_2) \psi_{c_2}(\mathbf{r}_1) \right]. \end{aligned}$$

Разность двух членов указывает на два различных процесса, изображенных на рис. 21. При этом состояние $|\alpha\rangle$ резонансно взаимодействует с такими состояниями $|k\rangle$, энергия которых приближенно равна $E_k \approx E_{c_2} + E_{c_3} - E_{c_1}$. Обратим внимание на то, что для точного удовлетворения данного соотношения по крайней мере одно состояние, участвующее в процессе, должно принадлежать континууму.

Пренебрежем непосредственным возбуждением одним фотоном двухэлектронного состояния $|k\rangle$ (вероятность этого процесса обычно мала). Будем считать, что для рассматриваемых состояний только матричный элемент $M_{\alpha g} \neq 0$ отличен от нуля. При этом для расчета мнимой части диэлектрической проницаемости необходимо вычислить только диагональный матричный элемент функции Грина $\mathcal{G}_{\alpha\alpha}$, поскольку

$$\varepsilon_2(\omega) \sim \pi^{-1} \text{Im}(M_{g\alpha} \mathcal{G}_{\alpha\alpha} M_{\alpha g}).$$

Для $\mathcal{G}_{\alpha\alpha}$ из (7.5) получаем уравнение

$$\begin{aligned} \mathcal{G}_{\alpha\alpha} &= \mathcal{G}_{\alpha\alpha}^0 + \mathcal{G}_{\alpha\alpha}^0 \mathcal{V}_{\alpha\alpha} \mathcal{G}_{\alpha\alpha}^0 + \sum_{\beta} \mathcal{G}_{\alpha\alpha}^0 \mathcal{V}_{\alpha\beta} \mathcal{G}_{\beta\beta}^0 \mathcal{V}_{\beta\alpha} \mathcal{G}_{\alpha\alpha}^0 + \\ &+ \sum_{\beta\gamma} \mathcal{G}_{\alpha\alpha}^0 \mathcal{V}_{\alpha\beta} \mathcal{G}_{\beta\beta}^0 \mathcal{V}_{\beta\gamma} \mathcal{G}_{\gamma\gamma}^0 \mathcal{V}_{\gamma\alpha} \mathcal{G}_{\alpha\alpha}^0 + \dots \end{aligned} \quad (7.6)$$

Здесь учтено, что $\mathcal{G}_{\alpha\beta}^0 = \delta_{\alpha\beta} / (\hbar\omega - i0 - E_{\alpha})$ диагональна по индексам α, β . Мы будем предполагать, что состояние $|\alpha\rangle$ является состоянием Хартри–Фока, т. е. корреляционные и

обменные эффекты включены в энергию E_α . Тогда $\mathcal{V}_{\alpha\alpha} = 0$, и второй член в (7.6) исчезает. В третьем члене в качестве состояния β необходимо использовать состояния $|k\rangle$. В этом случае он приобретает вид

$$\sum_k \frac{|\mathcal{V}_{k\alpha}|^2}{\hbar\omega - E_k - i0} \frac{1}{(\hbar\omega - E_\alpha - i0)^2}.$$

Будем считать, что взаимодействие между различными состояниями $|k\rangle$ и $|k'\rangle$ мало: $|\mathcal{V}_{kk'}| \ll |\mathcal{V}_{k\alpha}|$. Тогда во всех членах ряда необходимо оставить лишь суммы, содержащие $|\mathcal{V}_{k\alpha}|^2$, поскольку четные члены ряда будут содержать малые $|\mathcal{V}_{kk'}|$. Так, следующий неисчезающий член имеет вид

$$\begin{aligned} \sum_{kk'} \frac{|\mathcal{V}_{k\alpha}|^2}{\hbar\omega - E_k - i0} \frac{|\mathcal{V}_{k'\alpha}|^2}{\hbar\omega - E_{k'} - i0} \frac{1}{(\hbar\omega - E_\alpha - i0)^3} = \\ = \frac{1}{(\hbar\omega - E_\alpha - i0)^3} \left[\sum_k \frac{|\mathcal{V}_{k\alpha}|^2}{\hbar\omega - E_k - i0} \right]^2. \end{aligned}$$

Этот процесс можно продолжить далее, и мы получим

$$\begin{aligned} \mathcal{G}_{\alpha\alpha} &= \frac{1}{\hbar\omega - E_\alpha - i0} \left[1 + \frac{\delta E_\alpha(\omega)}{\hbar\omega - E_\alpha - i0} + \left(\frac{\delta E_\alpha(\omega)}{\hbar\omega - E_\alpha - i0} \right)^2 + \dots \right] = \\ &= \frac{1}{\hbar\omega - E_\alpha - i0} \left[1 - \frac{\delta E_\alpha(\omega)}{\hbar\omega - E_\alpha - i0} \right]^{-1} = \frac{1}{\hbar\omega - E_\alpha - \delta E_\alpha(\omega)}, \end{aligned}$$

где была использована формула геометрической прогрессии, и введены обозначения

$$\delta E_\alpha(\omega) \equiv \Delta_\alpha(\omega) + i\Gamma_\alpha(\omega) \equiv \sum_k \frac{|\mathcal{V}_{k\alpha}|^2}{\hbar\omega - E_k - i0}. \quad (7.7)$$

Тем самым взаимодействие приводит к эффективному сдвигу уровня Δ_α и к его уширению Γ_α . В случае непрерывного спектра суммирование можно заменить интегрированием по k с соответствующей заменой нормировки волновых функций:

$$\Delta_\alpha(\omega) \equiv V \cdot p \cdot \int dk \frac{|\mathcal{V}_{k\alpha}|^2}{\hbar\omega - E_k}, \quad \Gamma_\alpha(\omega) \equiv \pi \int dk |\mathcal{V}_{k\alpha}|^2 \delta(\hbar\omega - E_k).$$

Из этих формул видно, что сдвиг энергии перехода Δ_α определяется виртуальными процессами, а ширина уровня и, соответственно, время жизни $\tau_\alpha = \pi\hbar/\Gamma_\alpha$ — реальными переходами в состояния непрерывного спектра $|k\rangle$ (сравните с аналогичными рассуждениями в §3).

Мнимая часть диэлектрической проницаемости принимает окончательную форму

$$\varepsilon_2(\omega) \sim \frac{1}{\pi} \frac{|\mathcal{M}_{g\alpha}|^2 \Gamma_\alpha(\omega)}{(\hbar\omega - E_\alpha - \Delta_\alpha(\omega))^2 + \Gamma_\alpha^2(\omega)}.$$

Если пользоваться традиционной теорией возмущения, то поправка к сдвигу уровня и его ширина будут описываться выражениями типа (7.7), но не зависящими от частоты ω внешнего поля:

$$\delta E_\alpha^{(2)}(\omega) \equiv \Delta_\alpha^{(2)}(\omega) + i\Gamma_\alpha^{(2)}(\omega) \equiv \sum_k \frac{|\mathcal{V}_{k\alpha}|^2}{E_\alpha - E_k - i0}.$$

Если состояние $|k\rangle$ находится не около точки экстремума, например, дна зоны проводимости, то $\delta E_\alpha(\omega) \approx \delta E_\alpha^{(2)}(\omega)$ и слабо зависит от частоты ω . Линия поглощения при этом представляет из себя симметричный лоренциан с шириной Γ_α/\hbar . Если состояние $|k\rangle$ оказывается рядом с экстремумом зоны проводимости, линия становится асимметричной, и зависимостью Δ_α и Γ_α от частоты пренебрегать нельзя. Если $\Gamma_\alpha(\omega)$ растет с ростом частоты, то линия поглощения резко обрывается со стороны низких энергий и растягивается со стороны больших энергий.

Эффект Оже играет большую роль в распаде высокоэнергетических состояний, деля энергию на мелкие порции. Вероятность обратного процесса — переход из двухэлектронного состояния в одноэлектронное — обычно намного меньше, поэтому процесс идет только в направлении размена энергии. Это связано с тем, что число двухэлектронных состояний намного больше числа одноэлектронных состояний. Кроме того, состояния в зоне проводимости взаимодействуют друг с другом и с

другими квазичастицами, например, с фононами, и быстро выходят из резонанса с начальным одноэлектронным состоянием. Ниже мы рассмотрим процесс релаксации к основному состоянию более подробно, в частности, на примере систем, в которых возможно дробление энергии на малые порции — испускание фононов.

Здесь был рассмотрен только процесс уширения глубоких состояний в твердом теле за счет взаимодействия с продольным полем. Взаимодействие с поперечным электромагнитным полем также приводит к уширению линии — за счет конечной вероятности испускания поперечного фотона. Тем самым ширина линии Γ_α представляется суммой вкладов различных процессов: $\Gamma_\alpha = \Gamma_\alpha^{\text{Aug}} + \Gamma_\alpha^{\text{rad}}$. Радиационная ширина $\Gamma_\alpha^{\text{rad}}$ сильно зависит от частоты испускаемого фотона:

$$\Gamma_\alpha^{\text{rad}} = \hbar\tau_r^{-1}/2 = \frac{2\omega^3 |\mathbf{d}_{fi}|^2}{3c^3}.$$

(мы использовали известную формулу для радиационного времени жизни для двухуровневой системы; эта формула будет также получена в четвертой части). Для большинства не очень глубоких состояний (то есть за исключением K -состояний атомов с зарядом $Z > 40$) радиационное уширение мало по сравнению с Оже-уширением, и только для тяжелых элементов с увеличением разности энергий между уровнями (и тем самым с ростом ω) вероятность излучательных процессов становится значительной.

7.3 Эффект Фано

В рассмотренном выше эффекте Оже не учитывалось прямое взаимодействие света с конкурирующим процессом, поскольку поглощение с одновременным возбуждением двух электронов имеет очень малую вероятность. Но если схему, показанную на рис. 21, немного изменить, например, рассмотреть состояния c_3 и α как соответствующие одному и тому же уровню, то в результате Оже-процесса возникает состояние не с двумя, а только с одним возбужденным электроном. Такое состояние может быть получено и прямым возбуждением электрона с уровня c_2 , что приводит к конкуренции двух различных процессов поглощения света. Один из этих процессов (переход с уровня c_1 на уровень α) является

резонансным, то есть при пренебрежении взаимодействием соответствующая линия имеет нулевую ширину, а другой (переход с уровня c_2 в состояние непрерывного спектра k) — нерезонансный. Эффект Фано [84] как раз и описывает появление узкой линии поглощения на фоне непрерывного поглощения. Ранее при описании процесса поглощения света, например в модели невзаимодействующих электронов, предполагалось, что полная вероятность поглощения является суммой парциальных вероятностей отдельных процессов. Это перестает быть справедливым, если различные состояния взаимодействуют между собой.

На рис. 22 приведена схема уровней, иллюстрирующая эффект Фано. Под действием фотона с энергией $\hbar\omega$ возможны переходы из основного состояния $|g\rangle$ как в состояние $|\alpha\rangle = a_\alpha^+ a_{c_1} |g\rangle$, так и в состояние $|l\rangle = a_k^+ a_{c_2} |g\rangle$. Таким образом, отличны от нуля матричные элементы дипольных переходов $M_{g\alpha}$ и M_{gl} . Однако состояние $|\alpha\rangle$ взаимодействует с состоянием $|l\rangle$ через электростатическое поле. Матричный элемент оператора взаимодействия имеет вид

$$\begin{aligned} \mathcal{V}_{\alpha l} &= \left\langle \alpha \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| l \right\rangle = \left\langle g \left| a_{c_2}^+ a_k \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} a_\alpha^+ a_{c_1} \right| g \right\rangle = \\ &= \int d^3 r_1 d^3 r_2 \psi_{c_1}^*(\mathbf{r}_1) \psi_l^*(\mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \times \\ &\times \left[\psi_{c_2}(\mathbf{r}_2) \psi_\alpha(\mathbf{r}_1) - \psi_\alpha(\mathbf{r}_2) \psi_{c_2}(\mathbf{r}_1) \right]. \end{aligned}$$

Первый член соответствует процессу, показанному на рис. 22б, а второй — на рис. 22а. Знак минус возникает из-за фермионной статистики электронов. В этих предположениях форма линии поглощения принимает следующий вид:

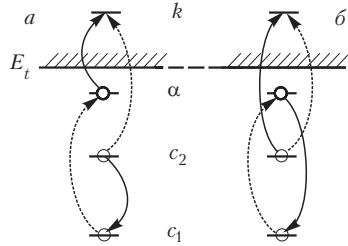


Рис. 22. Схемы переходов при описании эффекта Фано. Радиационные переходы показаны пунктирными стрелками.

$$\varepsilon_2(\omega) \sim \pi^{-1} \operatorname{Im} [M_{g\alpha} \mathcal{G}_{\alpha\alpha} M_{\alpha g} + \\ + \sum_l (M_{gl} \mathcal{G}_{l\alpha} M_{\alpha g} + M_{g\alpha} \mathcal{G}_{\alpha l} M_{lg}) + \sum_{l'l'} M_{gl} \mathcal{G}_{l'l'} M_{l'g}].$$

Здесь матричный элемент функции Грина $\mathcal{G}_{\alpha\alpha}$ учитывает взаимодействие со всеми состояниями l по аналогии с эффектом Оже:

$$\mathcal{G}_{\alpha\alpha} = (\hbar\omega - E_\alpha - \Delta_\alpha(\omega) - i\Gamma_\alpha(\omega))^{-1},$$

где $\Delta_\alpha(\omega) + i\Gamma_\alpha(\omega) \equiv \sum_l \frac{|\mathcal{V}_{kl}|^2}{\hbar\omega - E_l - i0}$, а остальные матричные

элементы функции Грина можно выразить с помощью уравнения Дайсона:

$$\mathcal{G}_{\alpha l} = \frac{\mathcal{G}_{\alpha\alpha} \mathcal{V}_{\alpha l}}{\hbar\omega - E_l - i0}; \quad \mathcal{G}_{l\alpha} = \frac{\mathcal{G}_{\alpha\alpha} \mathcal{V}_{l\alpha}}{\hbar\omega - E_l - i0}; \\ \mathcal{G}_{l'l'} = \frac{\delta_{l'l'}}{\hbar\omega - E_l - i0} + \frac{\mathcal{V}_{l\alpha} \mathcal{G}_{\alpha\alpha} \mathcal{V}_{\alpha l'}}{(\hbar\omega - E_l - i0)(\hbar\omega - E_{l'} - i0)}.$$

В этих формулах учтены лишь первые исчезающие поправки к затравочной функции Грина. Во многих случаях можно считать волновые функции действительными (например, для невырожденных состояний без магнитного поля), и тогда $\mathcal{V}_{l\alpha} = \mathcal{V}_{\alpha l}$ и $M_{\alpha g} = M_{g\alpha}$, $M_{lg} = M_{gl}$. В этом случае после достаточно громоздких, но простых алгебраических преобразований получим формулу Фано:

$$\varepsilon_2(\omega) = I_0 \frac{(q + \varepsilon')}{1 + \varepsilon'^2},$$

где $I_0 = \sum_l |M_{gl}|^2 \delta(E_\alpha - E_l)$ — интенсивность поглощения континуумом в отсутствие дискретного состояния $|\alpha\rangle$ при $\hbar\omega = E_\alpha$, а

$$\varepsilon' = (\hbar\omega - E_\alpha - \Delta_\alpha(E_\alpha)) / \Gamma_\alpha(E_\alpha)$$

— безразмерная энергия фотона, учитывающая ширину состояния $|\alpha\rangle$ и сдвиг энергии за счет взаимодействия Оже. Параметр q , от которого зависит форма линии поглощения, определяется интерференцией переходов в различные состояния:

$$q = \frac{M_{g\alpha} + \text{V.p.} \sum_l M_{gl} \mathcal{V}_{lg} / (E_\alpha - E_l)}{\pi \sum_l M_{gl} \mathcal{V}_{lg} \delta(E_\alpha - E_l)}.$$

При выводе этих формул учтено, что состояния не вырождены, состояние $|l\rangle$ находится далеко от порога поглощения, так что зависимостью I_0 , Δ_α и Γ_α от ω можно пренебречь.

В предельном случае отсутствия взаимодействия ($\mathcal{V}_{l\alpha} \rightarrow 0$) имеем $\Gamma_\alpha \rightarrow 0$, $q \rightarrow \infty$, и линия поглощения приобретает вид

$$\varepsilon_2(\omega) = I_0 + |M_{\alpha g}|^2 \delta(\hbar\omega - E_\alpha),$$

то есть имеет вид суммы двух независимых вкладов от непрерывного и дискретного состояний. В случае сильного взаимодействия линия квазидискретного поглощения расширяется, и ее форма становится асимметричной (рис. 23) (см., например, [85]). В случае $q=0$ вместо линии поглощения в спектре наблюдается провал, который называется антирезонансом. Таким образом, взаимодействие между состояниями приводит при некоторых энергиях не к усилению поглощения, а к его значительному ослаблению. Такой провал наблюдается и при других значениях q .

Поскольку помимо рассмотренной пары состояний существуют еще состояния, принимающие участие в поглощении фотона с энергией $\hbar\omega$, но слабее взаимодействующие с дискретным уровнем, провал происходит не до нуля.

Наиболее яркое экспериментальное проявление эффекта Фано обнаружено в твердом аргоне Ag в области основного поглощения с 3s-состояний Ar (26–29 эВ) [86]. При этом

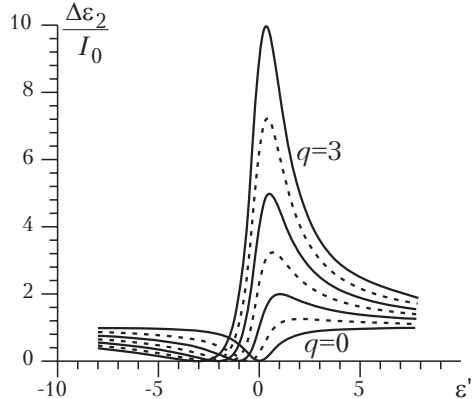


Рис. 23. Форма линии поглощения для эффекта Фано.

переходы $3s \rightarrow np$ экситонного типа наложены на непрерывный спектр поглощения, связанный с переходами из валентной зоны, образованной в основном $3p$ -состояниями Ag, в зону проводимости (рис. 24). Аналогичные формы линии наблюдались и в ряде других случаев.

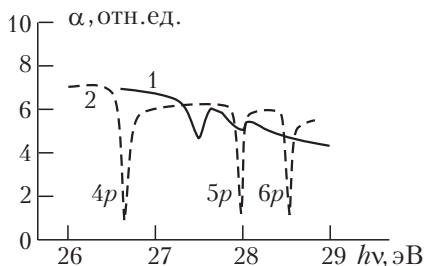


Рис. 24. Проявление эффекта Фано в твердом аргоне [86].

§8 Поляритоны

На примере взаимодействия света с экситонами рассмотрим еще одно интересное явление — возникновение поляритонов. В разделе 1.1 было показано, что уравнения для продольных и поперечных электромагнитных возбуждений имеют вид (1.13) и (1.14). При описании продольных и поперечных экситонов в разделе 6.2 указывалось, что энергия продольного экситона E_{L0} удовлетворяет уравнению (1.13), а энергия поперечного экситона E_{T0} — уравнению

$$\tilde{\epsilon}^{-1}(E_{T0}/\hbar) = 0. \quad (8.1)$$

Это уравнение отличается от уравнения (1.14) и переходит в него лишь в пределе $c \rightarrow \infty$. Решения уравнений для продольного и поперечного экситонов не зависят от волнового вектора экситона \mathbf{k} (поскольку диэлектрическая проницаемость (6.7) не зависит от \mathbf{k}) и показано на рис. 25 пунктирной линией 2 и сплошной линией 3, параллельными оси k .

Условие $c \rightarrow \infty$ соответствует пренебрежению запаздывающим взаимодействием с фотонами, поскольку при рассмотрении экситонов предполагается, что электромагнитное поле имеет кулоновский характер. Решение точного уравнения (1.14) зависит от \mathbf{k} даже если $\tilde{\epsilon}$ не зависит от \mathbf{k} . Более того, уравнение (1.14) имеет действительные решения для всех действительных ω , за исключением области $E_{T0} < \hbar\omega < E_{L0}$ (для этой

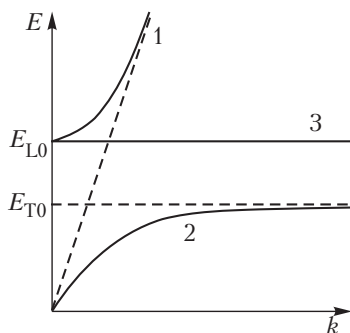


Рис. 25. Дисперсионные кривые для продольных (3) и поперечных (1, 2) возбуждений с учетом запаздывания.

области $\varepsilon_1(\omega) < 0$ и поэтому k получается чисто мнимым). Решение (1.14) может быть записано в виде

$$k^2 = c^{-2} \omega^2 \tilde{\varepsilon}(\omega) \quad (8.2)$$

и выходит на асимптотические линейные зависимости на малых и больших частотах:

$$k = \omega \sqrt{\varepsilon_0} / c \quad \text{для } \omega \ll \Omega_T \equiv E_{T0} / \hbar, \quad (8.3)$$

$$k = \omega \sqrt{\varepsilon_\infty} / c \quad \text{Для } \omega \gg \Omega_L \equiv E_{L0} / \hbar. \quad (8.4)$$

Прямая линия (8.4) показана на рис. 25 пунктирной линией 1, а точные решения (8.2) – сплошными линиями 1 и 2. Возбуждения с законом дисперсии (8.2) называются поляритонами. Поскольку эти возбуждения являются решениями точного уравнения (1.14) а поперечные экситоны – лишь приближенного уравнения (8.1), описание возбуждений на языке поперечных экситонов не полностью соответствует реальной картине.

Если учитывать пространственную дисперсию диэлектрической проницаемости, дисперсионные кривые для поляритонов отличаются от показанных на рис. 25. Если эффективная масса экситона положительна, горизонтальные линии 2 и 3 превращаются в параболы (причем энергия E растет с ростом k). Пространственная дисперсия придает поляритонам новые черты: во-первых, появляются действительные решения для поперечных возбуждений в области $\Omega_T < \omega < \Omega_L$, для которой в отсутствие пространственной дисперсии не было решений, и, во-вторых, в области $\omega > \Omega_L$ возникает два решения, одно из которых соответствует большим k , а другое – малым k . Если эффективная масса экситона отрицательна, горизонтальные линии изгибаются вниз. При этом область отсутствия действительных решений расширяется в сторону малых энергий, и два решения возникают уже ниже Ω_T , а не выше Ω_L .

Возникновение двух решений с одним и тем же значением ω требует дополнительных граничных условий для задачи преломления света, падающего из вакуума в кристалл. Такие дополнительные граничные условия могут быть получены только при использовании микроскопической теории, когда правильное поведение $\tilde{\varepsilon}(\omega, \mathbf{k})$ может быть получено для всех \mathbf{k} (см., например, [10]).

Приведем феноменологическое описание процесса отражения света, падающего по нормали на полубесконечный кристалл. Если пространственная дисперсия сильная, обычные формулы Френеля (1.18), (1.19) перестают быть справедливыми. В одном из подходов к решению этой задачи [1] предполагается, что поверхность кристалла можно смоделировать, введя поверхностный токовый слой, плотность тока в котором описывается законом

$$\mathbf{j}(\mathbf{r}, t) = \mathbf{j}_0 e^{-i\omega t} \delta(x)$$

(предполагается, что кристалл заполняет полупространство $x > 0$). Поверхностный ток \mathbf{j}_0 возбуждается внешним электромагнитным излучением. Электрическое и магнитное поле в кристалле ($x > 0$), индуцированное поверхностным током, может быть получено из (1.6) с помощью Фурье-преобразования:

$$E_y^>(x, t) = \frac{2}{i\omega} j_{0y} e^{-i\omega t} \int_{-\infty}^{\infty} dk \frac{e^{ikx}}{\tilde{\epsilon}(\omega, k) - c^2 k^2 / \omega^2}, \quad (8.5)$$

$$H_z^>(x, t) = \frac{2c}{i\omega^2} j_{0y} e^{-i\omega t} \int_{-\infty}^{\infty} dk \frac{ke^{ikx}}{\tilde{\epsilon}(\omega, k) - c^2 k^2 / \omega^2} \quad (8.6)$$

(предполагаем, что падающий свет поляризован в направлении y).

Поля в области $x < 0$ описываются суммой падающей и отраженной волны: $E_y^< = E_y^i + E_y^r$, $H_z^< = H_z^i + H_z^r = E_y^i - E_y^r$ (в последнем соотношении использовались свойства свободных волн в вакууме). Поскольку поперечные компоненты полей должны быть непрерывны, можно записать $E_y^<(0, t) = E_y^>(0, t)$, $H_z^<(0, t) = H_z^>(0, t)$. Тем самым коэффициент отражения $R = \left| E_y^r / E_y^i \right|^2$ может быть выражен через интегралы (8.5), (8.6):

$$R = \left| \frac{\tilde{Q} - 1}{\tilde{Q} + 1} \right|^2 \quad (8.7)$$

где

$$\tilde{Q} = \int_{-\infty}^{\infty} \frac{N dN}{\tilde{\varepsilon}(\omega, \omega N/c) - N^2} \bigg/ \int_{-\infty}^{\infty} \frac{dN}{\tilde{\varepsilon}(\omega, \omega N/c) - N^2} \quad (8.8)$$

(здесь k заменено на $N\omega/c$). Если диэлектрическая проницаемость не зависит от k , подынтегральные выражения в (8.8) имеют только один полюс $\tilde{N}_0(\omega) = \sqrt{\tilde{\varepsilon}(\omega)}$, и \tilde{Q} становится обычным показателем преломления, $\tilde{Q} = \tilde{N}_0$. Для обсужденной выше модели экситонов Френкеля (6.7) $\tilde{N}_0(\omega)$ действительно для $\omega < \Omega_T$ и $\omega > \Omega_L$ и мнимо в области $\Omega_T < \omega < \Omega_L$. Поэтому коэффициент отражения в последнем интервале частот обращается в единицу (в кристалле не могут существовать возбуждения с такими частотами).

Если пространственная дисперсия становится сильной, в подынтегральных выражениях может оказаться несколько полюсов. Каждый полюс \tilde{N}_i соответствует волне с волновым вектором $\tilde{N}_i\omega/c$. Вклад каждой из таких волн определяется вычетом \tilde{B}_i подынтегрального выражения:

$$\tilde{Q} = \sum_i \tilde{N}_i \tilde{B}_i / \sum_i \tilde{B}_i, \text{ где}$$

$$\tilde{B}_i = \left[\frac{\omega}{c} \frac{\partial \tilde{\varepsilon}(\omega, \tilde{k})}{\partial \tilde{k}} \bigg|_{\tilde{k} = \tilde{N}_i \omega/c} - 2\tilde{N}_i \right]^{-1},$$

\tilde{N}_i — решения уравнения $\tilde{\varepsilon}(\omega, \tilde{N}\omega/c) = \tilde{N}^2$ с положительной мнимой частью. Электромагнитное поле в кристалле также становится суперпозицией частных решений:

$$E_y^>(x, t) \sim \sum_i B_i e^{i\tilde{N}_i \omega x/c - i\omega t}.$$

Для положительных масс экситона в области между Ω_T и Ω_L имеется два решения для \tilde{N} , одно из которых чисто мнимое, а другое действительное. Тем самым коэффициент отражения в этой области уже не равен единице, поскольку могут возбуждаться экситоны.

Компонента волнового вектора вдоль оси x для создаваемых светом возбуждений не связана непосредственно с волновым вектором свободного фотона, поскольку граница снимает трансляционную инвариантность. Поэтому длинноволновый

свет может создавать и коротковолновые возбуждения. Естественно, что компоненты волнового вектора, параллельные поверхности, сохраняются.

При исследовании прохождения света через плоскопараллельную пластину необходимо учитывать отражение от задней грани и интерференцию волн. Условия интерференции отличаются для волн с различными волновыми векторами $\tilde{N}_i\omega/c$. Тем самым коэффициент прохождения света через пластину также может существенно изменяться. Многочисленные примеры приведены в [1, 10].

Квантовомеханическое рассмотрение задачи о поляритоне показывает, что поляритон является комбинацией экситона и фотона. Доля экситонного вклада в поляритонное состояние меняется при движении вдоль дисперсионной кривой. На участках кривых, близких к горизонтали (с относительно большими k , где на рис. 25 совпадают сплошная и штриховая линии), поляритон почти совпадает с экситоном — реальным возбуждением электронов и дырок. На наклонных участках, где дисперсионная кривая совпадает с асимптотиками (8.3), (8.4), поляритон представляет собой фотон с малой примесью виртуально возбужденных экситонов.

Наиболее интересной частью кривой является переходный участок, на котором поляритон часть времени как бы находится в состоянии фотона, а часть — в состоянии экситона. Отклонения от чистых состояний и величина переходной области увеличиваются с ростом силы осциллятора на частоте Ω_T , то есть с ростом разности $\epsilon_0 - \epsilon_\infty$. Поэтому учитывать существование поляритонов имеет смысл только в случае больших сил осциллятора (например, в случае, когда коэффициент поглощения в экситонной полосе составляет 10^6 см^{-1} , что характерно для некоторых молекулярных кристаллов и частично для твердых инертных газов).

Поляритонная модель может быть применена для объяснения экситонной люминесценции. Для этого необходимо ввести взаимодействие с фононами или другими квазичастицами. При учете такого взаимодействия поляритоны могут передавать импульс фононам, поглощая и испуская их, и тем самым передвигаться по дисперсионной кривой. Казалось бы, это должно привести к быстрому скатыванию поперечного поляритона нижней ветви в состояние с нулевой энергией, поскольку эта поляритонная ветвь плавно уходит в точку $\omega = 0$, $\mathbf{k} = 0$. Однако с фононами эффективно взаимодействует

только экситонная часть поляритона, а фотонная не взаимодействует с фононами. Поэтому бoльцмановское распределение поляритонов по энергии не достигается, а поляритоны скапливаются в области выше резкого изгиба дисперсионной кривой. Этот эффект называется «эффектом бутылочного горла». Тем самым выход фотонов из кристалла (переход поляритона через границу кристалла) происходит на частотах, лишь немногие ниже Ω_T . Сдвиг и ширина линии люминесценции возрастает с ростом силы осциллятора. С ростом силы осциллятора меняется также и пространственное распределение поляритонов.

Можно показать, что время жизни поляритона, т. е. время прохода поляритона через узкое горло, оказывается равным при таком рассмотрении времени радиационного высвечивания, полученного по теории возмущения (для умеренных значений сил осциллятора). Тем самым устанавливается принцип соответствия между рассмотрением взаимодействия со светом по теории возмущения и поляритонным описанием этого явления.

§9 Электронные состояния дефектов в недеформируемой решетке

9.1 Классификация дефектов в твердых телах

Выше были рассмотрены эффекты, приводящие к поглощению фотона в идеальном недеформируемом кристалле. Однако реальные кристаллы содержат большое число дефектов и примесей, которые образуют свои электронные состояния. Несмотря на то, что концентрация такого рода примесей может быть малой, их роль в спектроскопии твердого тела оказывается значительной, особенно при изучении эффектов, связанных с поглощением и излучением фотонов, энергия которых ниже порога фундаментального поглощения. Дефектам в твердых телах посвящена обширная литература (см., например, [5]). Изучению такого рода дефектов посвящены следующие параграфы.

Число различных типов дефектов чрезвычайно велико, поэтому даже вопрос классификации дефектов не может быть решен однозначно. Его решение зависит часто от того, какие наблюдаемые свойства дефектов интересуют исследователя.

Если интересоваться вопросом происхождения дефектов, то их можно разделить на собственные и примесные дефекты. Понятие примесного дефекта очевидно — это дефекты, обусловленные атомами, не входящими в состав основного вещества. Собственными дефектами могут являться вакансии, атомы в междоузлиях и т. д.

Геометрические свойства дефектов позволяют разделить их на точечные, связанные с отдельными атомами, линейные (например, дислокации) и плоские (например, плоскость скольжения), объемные (включения и поры). В принципе границу кристалла также можно отнести к разряду плоских дефектов, поскольку, во-первых, на границе нарушается периодичность кристалла, во-вторых, часто на границе меняется структура кристалла.

Дефекты могут захватывать или отдавать электроны. В теории полупроводников такого рода дефекты называются акцепторами или донорами. К обоим типам дефектов применимо название ловушки — электронной или дырочной в зависимости от того, какой процесс захвата является основным.

Дефекты классифицируются также по отношению к взаимодействию со светом. Так, дефекты, имеющие такие электронные состояния, энергия переходов между которыми лежит в области запрещенной зоны, называются центрами окраски, поскольку именно они приводят к окраске прозрачных в видимой области кристаллов. Если дефект интенсивно захватывает электронные возбуждения и приводит к излучательной рекомбинации в виде фотона люминесценции, то такого рода дефекты называются центром свечения.

Такого рода классификацию можно продолжать и далее. Необходимо только отметить условность многих названий дефектов, поскольку в различных условиях одни и те же дефекты могут вести себя по-разному.

В твердом теле любой дефект описывается изменением некоторых членов в гамильтониане идеального кристалла. При этом сразу нарушается одно из основных преимуществ в описании кристалла — его трансляционная инвариантность. Функции электронов в кристалле уже перестают удовлетворять теореме Блоха, и волновой вектор перестает быть «хорошим» квантовым числом. Волновые функции идеального кристалла используются только в качестве начального приближения.

В отношении дефектов можно выделить те же два подхода, какие встречались при описании экситонов. Начнем с описания дефектов большого радиуса.

9.2 Дефекты большого радиуса

Если предположить, что возмущение, которое вносит потенциал дефекта, является слабым, то можно использовать в качестве функций нулевого приближения блоховские функции идеального кристалла. На этом базисе можно построить волновую функцию дефекта (см., например, [87]):

$$|\lambda\rangle = \sum_{\alpha\mathbf{k}} f_{\lambda\alpha}(\mathbf{k}) |\alpha\mathbf{k}\rangle, \quad (9.1)$$

где $|\alpha\mathbf{k}\rangle = V^{-1/2} e^{i\mathbf{k}\mathbf{r}} u_{\alpha\mathbf{k}}(\mathbf{r})$ — блоховская функция (7.4)

электрона в кристалле, принадлежащего зоне α . Волновые функции $|\alpha\mathbf{k}\rangle$ предполагаются удовлетворяющими уравнению идеального кристалла:

$$\hat{\mathcal{H}}_0|\alpha\mathbf{k}\rangle = E_\alpha(\mathbf{k})|\alpha\mathbf{k}\rangle.$$

Дефект вносит возмущение $\hat{\mathcal{V}} = \hat{\mathcal{H}}_1 - \hat{\mathcal{H}}_0$, где $\hat{\mathcal{H}}_1$ — гамильтониан кристалла с дефектом. Тем самым для определения энергии и волновой функции состояния необходимо решить уравнение

$$(\hat{\mathcal{H}}_0 + \hat{\mathcal{V}})|\lambda\rangle = E_\lambda|\lambda\rangle.$$

Домножая это уравнение слева на $\langle\alpha\mathbf{k}|$ и интегрируя по всему пространству, получим

$$[E_\alpha(\mathbf{k}) - E_\lambda]f_{\lambda\alpha}(\mathbf{k}) + \sum_{\alpha\mathbf{k}'} \int d^3r \langle\alpha\mathbf{k}|\hat{\mathcal{V}}(\mathbf{r})|\alpha'\mathbf{k}'\rangle f_{\lambda\alpha'}(\mathbf{k}') = 0.$$

Перейдем в этом уравнении к огибающим $F_{\lambda\alpha}(\mathbf{l}) = N^{-1/2} \sum_{\mathbf{k}} e^{i\mathbf{k}\mathbf{l}} |\alpha\mathbf{l}\rangle$. Предварительно преобразуем входящий в это уравнение интеграл, перейдя от функций Блоха к функциям Ванье, локализованными вблизи ячейки с радиус-вектором \mathbf{l} : $|\alpha\mathbf{l}\rangle = N^{-1/2} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{l}} |\alpha\mathbf{k}\rangle$. После этого под интегралом по всему пространству будет стоять

$$N^{-1} \sum_{\Pi'} \langle\alpha\mathbf{l}|\hat{\mathcal{V}}(\mathbf{r})|\alpha'\mathbf{l}'\rangle e^{-i\mathbf{k}\mathbf{l} + i\mathbf{k}\mathbf{l}'}$$

Слабость потенциала возмущения $\hat{\mathcal{V}}(\mathbf{r})$ предполагает, в частности, слабую зависимость $\hat{\mathcal{V}}$ от координат на расстояниях порядка постоянной решетки. За счет этого можно вынести $\hat{\mathcal{V}}(\mathbf{l})$ за знак матричного элемента и затем воспользоваться полнотой системы функций Ванье:

$$\int \langle\alpha\mathbf{l}|\alpha'\mathbf{l}'\rangle d^3r = \delta_{\alpha\alpha'} \delta_{\Pi\Pi'}.$$

Уравнение для огибающей $F_{\lambda\alpha}$ в этом случае будет иметь вид (с заменой $\mathbf{l} \rightarrow \mathbf{r}$ в континуальном приближении):

$$[E_\alpha(\hat{\mathbf{k}}) - E_\lambda] F_{\lambda\alpha}(\mathbf{r}) + \mathcal{V}(\mathbf{r}) F_{\lambda\alpha}(\mathbf{r}) = 0.$$

Здесь $\hat{\mathbf{k}} = -i\nabla$, а функция $E_\alpha(\hat{\mathbf{k}})$ понимается как функция от оператора: $E_\alpha(-i\nabla)$. Это уравнение определяет энергетические уровни дефекта в зоне относительно уровня экстремума функции $E_\alpha(\mathbf{k})$. Разлагая $E_\alpha(\mathbf{k})$ в ряд по степеням $\mathbf{k} - \mathbf{k}_0$, где \mathbf{k}_0 — точка зоны Бриллюэна, в которой достигается экстремум зоны α , в изотропном случае без вырождения получим уравнение Шредингера для электрона с массой m_α^* в поле $\hat{\mathcal{V}}(\mathbf{r})$:

$$\left(-\frac{\hbar^2}{2m_\alpha^*} \nabla^2 + \mathcal{V}(\mathbf{r}) - E_\lambda \right) F_{\alpha\lambda}(\mathbf{r}) = 0 \quad (9.2)$$

Заметим, что влияние периодического поля кристалла сводится лишь к замене массы электрона на эффективную массу m_α^* . Волновая функция электрона может быть получена с использованием функции Блоха $|\alpha\mathbf{k}\rangle$. Поскольку Фурье-образ $f_{\alpha\lambda}(\mathbf{k})$ отличен от нуля лишь в небольшой области зоны Бриллюэна, то в уравнении (9.1) можно положить $\mathbf{k} = \mathbf{k}_0$ в функции $|\alpha\mathbf{k}\rangle$, и волновая функция электрона дефекта становится равной

$$\psi(\mathbf{r}) = \int d^3k f_{\lambda\alpha}(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} |\alpha\mathbf{k}_0\rangle = F_{\lambda\alpha}(\mathbf{r}) |\alpha\mathbf{k}_0\rangle.$$

Таким образом, $F_{\lambda\alpha}(\mathbf{r})$ имеет смысл огибающей периодической функции Блоха $|\alpha\mathbf{k}_0\rangle$, осциллирующей с периодом решетки.

Более сложные уравнения получаются в случае вырожденных состояний, однако здесь эти уравнения обсуждать не будем. Как уже говорилось, этот метод применим лишь для дефектов, возмущающий потенциал которых медленно меняется в пространстве, а радиус электронных состояний дефектов во много раз больше постоянной решетки. Наиболее частым случаем является учет кулоновского поля примеси — $Ze^2/\epsilon r$, где Ze — заряд примеси, ϵ — диэлектрическая проницаемость. Подстановка такой потенциальной энергии в уравнение приводит к водородоподобному уравнению. Поэтому спектр энергий такой примеси является ридберговским: $E_n = E_\alpha(\mathbf{k}_0) - Ry^*/n^2$, где $Ry^* = m_\alpha^* e^4 / 2\epsilon^2 \hbar^2$. Радиус такого состояния $a^* = a_B \epsilon m / m_\alpha^*$. Теория эффективной массы применима только для кристаллов, для которых a^* много больше постоянной решетки a , поэтому кристалл должен иметь малую

эффективную массу электронов m_{α}^* и большое значение ε . В этом случае такого рода состояния имеют малую энергию, поскольку Ry^* оказывается малым.

Теория эффективной массы применима также к высоко-возбужденным состояниям таких заряженных дефектов, основное состояние которых имеет малый радиус. В случае отсутствия свободных носителей заряда экранировки кулоновского потенциала не происходит, поэтому в диэлектрических кристаллах заряженные примеси будут иметь бесконечное число уровней, сгущающихся при приближении к разрешенной зоне. Эта серия уровней аналогична экситонной серии. Такого рода состояния можно рассматривать либо как высоко-возбужденные состояния дефекта, либо как состояния экситона большого радиуса, локализованного около дефекта. Различие здесь только терминологическое, физика обоих процессов совершенно одинакова.

Если условия применимости модели эффективной массы выполняются для нескольких зон, дефект будет создавать связанные состояния вблизи от каждой зоны. Если ионизованный дефект заряжен положительно, связанные состояния образуются у дна зоны; если отрицательно, то около потолка (связанные состояния электрона на отрицательно заряженном центре возможны только в случае отрицательной эффективной массы m_{α}^*).

9.3 Глубокие дефекты

Как выше отмечалось, теория эффективной массы может описывать лишь мелкие уровни. Такого рода дефекты рассчитываются многочисленными методами, некоторые из которых применимы и для расчета идеальных кристаллов. Однако из-за невозможности использования трансляционной симметрии последовательные расчеты такого рода чрезвычайно сложны. Поэтому достаточно широко применяются различные квантовохимические методы численного расчета системы атомов, включающей дефектную область (кластер с дефектом). В таких расчетах важно правильно учесть влияние кулоновского поля оставшейся части кристалла. Точность такого рода вычислений относительно невысока, однако ряд качественных результатов оказывается чрезвычайно полезным. Для иллюстрации рассмотрим влияние кристаллического поля на примеси ионов редкоземельных металлов.

Исследованный выше случай соответствует слабому потенциалу примеси, когда потенциал примеси можно рассматривать как возмущение по отношению к кристаллическому потенциалу. Однако часто реализуется обратный предельный случай, когда, напротив, кристаллическое поле мало по сравнению с атомным полем примеси в том месте, где находится орбита электрона. Такое положение осуществляется, например, в примесях переходных металлов и редкоземельных примесях, в которых частично заполненные d - и f -орбиты эффективно экранируются внешними электронами. При этом электронные состояния таких ионов относительно слабо возмущаются полем остального кристалла. Поэтому редкоземельные ионы, например, ведут себя почти одинаково в различных матрицах.

В качестве нулевого приближения удобно взять волновые функции изолированного атома (иона). Эти волновые функции будут соответствовать $(2L + 1)$ -кратно вырожденным уровням. Поскольку рассматриваемые ионы содержат один или больше электронов с $l = 2$ или $l = 3$, многоэлектронная волновая функция может иметь большую кратность вырождения (полное L от 2 до 6), т. е. она многократно вырождена. При помещении такого иона в кристалл возмущающий кристаллический потенциал должен вызвать расщепление многократно вырожденных состояний. Если положение иона в кристалле несимметричное, то уровень распадается на $(2L + 1)$ подуровень. Однако часто в кристаллах кубической симметрии ион занимает одну из позиций иона самого кристалла (дефект замещения). В таком случае снятие вырождения неполное.

Для примера рассмотрим случай кристалла с решеткой NaCl. В таком кристалле точечная группа симметрии положения иона будет O_h , и в ней возможны однократно, двухкратно и трехкратно вырожденные состояния. Чтобы получить расщепление уровней, можно воспользоваться теоремой разложения и разложить приводимые $(2L + 1)$ -вырожденные представления группы вращений на неприводимые представления группы O_h [49]. При этом можно получить следующие правила перехода:

$$\begin{aligned}
L = 0 & \rightarrow A_{1g}(\Gamma_1); \\
L = 1 & \rightarrow T_{1u}(\Gamma_{15}); \\
L = 2 & \rightarrow T_{2g}(\Gamma'_{25}) + E_g(\Gamma_{12}); \\
L = 3 & \rightarrow T_{1u}(\Gamma_{15}) + T_{2u}(\Gamma_{25}) + A_{2u}(\Gamma'_2); \\
L = 4 & \rightarrow T_{2g}(\Gamma'_{25}) + T_{1g}(\Gamma'_{15}) + E_g(\Gamma_{12}) + A_{1g}(\Gamma_1); \\
L = 5 & \rightarrow T_{2u}(\Gamma_{25}) + 2T_{1u}(\Gamma_{15}) + E_u(\Gamma'_{12}); \\
L = 6 & \rightarrow 2T_{2g}(\Gamma'_{25}) + T_{1g}(\Gamma'_{15}) + E_g(\Gamma_{12}) + \\
& + A_{2g}(\Gamma_2) + A_{1g}(\Gamma_1);
\end{aligned} \tag{9.3}$$

Здесь приведены молекулярные обозначения неприводимых представлений и, в скобках, соответствующие им обозначения неприводимых представлений в центре зоны Бриллюэна кристалла с решеткой NaCl. Напомним, что A — однократно, E — двукратно и T — трехкратно вырожденные состояния. Индекс g или u указывает на четность или нечетность соответствующих волновых функций. Тем самым P -состояния в кубическом поле не расщепляются, D -состояния расщепляются на два уровня и т. д. Таким образом можно определить число уровней, образуемых примесью.

Теория групп помогает и в вычислении смещений этих уровней относительно их положения в изолированном атоме. Для этого необходимо вычислить смещение энергии по теории возмущений:

$$\Delta E^{(\alpha)} = \int d^3r \psi^{(\alpha)*} V_{\text{Coul}} \psi^{(\alpha)},$$

где $\psi^{(\alpha)}$ — одна из волновых функций неприводимого представления α , а V_{Coul} — поле внешнего (по отношению к иону) кубического окружения. Вблизи от начала координат, где помещен ион примеси, V_{Coul} можно разложить в ряд по мультиполям:

$$V_{\text{Coul}}(\mathbf{r}) = e^2 \sum_{lm} A_l \alpha_l^m Y_l^m(\theta, \varphi) \frac{r^l}{a^{l+1}},$$

где a — постоянная решетки, Y_l^m — сферическая функция, зависящая от полярных координат вектора \mathbf{r} , а коэффициенты A_l и α_l^m можно вычислить, исходя из симметричных

соображений. В случае чисто ионного кристалла A_l определяется суммами по решетке типа сумм, определяющих постоянную Маделунга, и могут быть вычислены для каждого типа решетке.

Поскольку в $\Delta E^{(\alpha)}$ входят волновые функции одного того же представления α , то для вычисления $\Delta E^{(\alpha)}$ необходимо знать только полносимметричную часть V_{Coul} , принадлежащую представлению A_{1g} (Γ_1). Если суммирование по l ограничить $l \leq 7$, то в соответствии с (9.3) остаются только три члена в $\Delta E^{(\alpha)}$ (поскольку A_{1g} не содержится в разложении других состояний): $l = 0, l = 4, l = 6$. Тем самым для определения смещения уровней необходимо знать только зависящие от примеси интегралы $\langle \psi | r^4 | \psi \rangle$ и $\langle \psi | r^6 | \psi \rangle$, остальные параметры определяются типом решетке. Так, для d -электрона в кубическом поле можно получить

$$\Delta E_{T_{2g}} = -4Dq + \Delta E_0, \quad \Delta E_{E_g} = 6Dq + \Delta E_0,$$

где $q = e \langle r^4 \rangle / 105$, а $D = -eA_4 / a^5$. Для решетке типа NaCl $A_4 = 3,58$, для решетке типа CaF₂ $A_4 = 2,712$. Величина ΔE_0 определяет смещение центра тяжести двух уровней и определяется константой Маделунга для рассматриваемого кристалла. Из эксперимента можно оценить величину расщепления между уровнями, равную $10Dq$.

Учет спин-орбитального взаимодействия можно провести двумя путями в зависимости от силы спин-орбитального взаимодействия по сравнению с кристаллическим полем. Если кристаллическое поле вызывает более сильное расщепление, то спин-орбитальное взаимодействие можно учитывать в последнюю очередь (рис. 26, а). Если же, наоборот, спин-орбитальное взаимодействие сильнее кристаллического поля, необходимо вначале перейти к уровням с определенным значением полного момента J и лишь затем раскладывать представления двойной группы вращений с полуцелым J по представлениям двойной группы кристалла (рис. 26, б). Число и классификация уровней не зависят от порядка разложения, однако их расположение по энергии меняется. Первый случай реализуется в основном для примесей переходных металлов, а второй – для редкоземельных элементов, в которых f -электроны сильнее локализованы вблизи ядра.

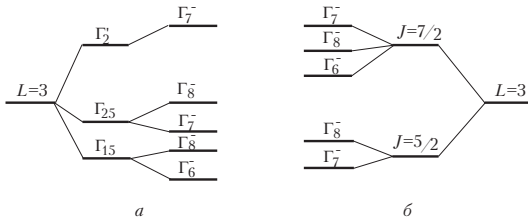


Рис. 26. Различные схемы образования уровней под действием кристаллического поля и спин-орбитального взаимодействия.

9.4 Плотность состояний в кристалле с дефектами

Глубокие дефектные состояния проявляются в плотности состояний в виде острых пиков. Положения этих пиков определяются энергиями соответствующих уровней дефектов, а интенсивность – концентрацией дефектов с учетом вырождения уровней.

Только центры, взаимодействующие с носителями заряда за счет кулоновского поля, могут образовывать квазинепрерывную систему уровней. Ограничимся именно этим случаем. Верхние уровни, приближающиеся к континууму, могут быть описаны как ридберговские состояния: $E = -Ry^*/n^2$, где n – главное квантовое число (см. раздел 6.2; здесь энергия отсчитывается от дна зоны). Каждый уровень энергии с определенным n вырожден n^2 раз по орбитальному и магнитному квантовому числу. Разность энергий между двумя последовательными состояниями равна $dE/dn = 2Ry^*/n^3$. Тем самым плотность состояний для высоковозбужденных состояний центров с концентрацией n_c равна

$$\begin{aligned}
 g(E) &= n_c V n^2 \frac{dn}{dE} = n_c V \frac{n^5}{2Ry^*} = \\
 &= n_c V \frac{Ry^*{}^{3/2}}{2|E|^{5/2}} = \frac{m^*{}^{3/2} e^6 n_c V}{2^{5/2} \epsilon^3 \hbar^3 |E|^{5/2}}
 \end{aligned}
 \tag{9.4}$$

Это выражение может быть выведено и с использованием квазиклассического приближения [54] из вычисления объема фазового пространства, соответствующего движению частицы

в притягивающем кулоновском потенциале. Для положительных энергий плотность состояний $g_f(E)$ для свободных частиц с кинетической энергией E дается уравнением (7.10). Кинетическая энергия частицы в кулоновском поле зависит от расстояния r до центра: $E_{\text{kin}} = E + e^2/\epsilon r$. Тем самым плотность состояний для $E < 0$ может быть получена интегрированием $g_f(E_{\text{kin}})$ по координатам в области, где потенциальная энергия меньше E :

$$g(E) = n_c \int d^3 r g_f(E + e^2/\epsilon r) \theta(E + e^2/\epsilon r) =$$

$$= \frac{4\pi n_c e^6}{\epsilon^3} \int_0^\infty \frac{dE_{\text{kin}}}{(E_{\text{kin}} + |E|)^4} g_f(E_{\text{kin}}). \quad (9.5)$$

Интегрирование этого выражения с подстановкой g_f из (7.10) опять дает формулу (9.4).

Поскольку концентрация n_c предполагается малой, наличие центров относительно слабо изменяет плотность состояний для положительных энергий. Тем самым $g(E) = g_f(E)$ для $E > 0$. При приближении к порогу снизу $|E| \rightarrow 0$ плотность $g(E)$ возрастает до бесконечности согласно (9.4), в то время как для положительных энергий $E \rightarrow 0$ плотность $g(E)$ падает до нуля согласно (4.10). Это противоречие может быть легко разрешено, если учесть, что вблизи от точки $E = 0$ ни уравнение (9.4), ни (4.10) не применимы. Условия их применимости нарушаются, поскольку для малых энергий волновая функция определяется одновременным взаимодействием с многими центрами и с самим кристаллом. Если потребовать, чтобы одно выражение плавно переходило в другое, то можно разложить точную плотность состояний в ряд вблизи точки $E = 0$: $g(E) \approx g(0) + Eg'(0)$, и это разложение должно плавно переходить в выражения (9.4) при энергии $-E_1$ и в (4.10) при энергии E_2 (см. рис. 27). Это

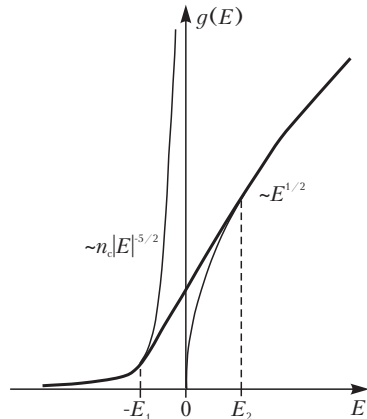


Рис. 27. Плотность состояния кристалла с заряженными дефектами.

дает

$$E_2 = \frac{7}{5} |E_1| = \frac{35^{1/6} \pi^{2/3}}{2^{2/3}} \frac{e^2 n_c^{1/3}}{\varepsilon} \approx 2,4 \frac{e^2 n_c^{1/3}}{\varepsilon}.$$

Энергия $e^2 n_c^{1/3} / \varepsilon$ равна средней энергии взаимодействия электрона с ближайшим центром в среде с диэлектрической проницаемостью ε , поскольку среднее расстояние от электрона до какого-либо центра равно $n_c^{-1/3}$. Значение плотности состояний около $E = 0$ определяется формулами

$$g(0) = \frac{m^{*3/2} V E_2^{1/2}}{2^{3/2} \pi^2 \hbar^3}, \quad g'(0) = g(0)/E_2. \quad (9.6)$$

Такой подход представляет из себя только грубое приближение, поскольку точное поведение плотности состояний в этой области должно изучаться методами теории неупорядоченных систем. Полное число состояний с отрицательной энергией может быть получено интегрированием уравнения (9.4) по энергии от минус бесконечности до $-E_1$:

$$\int_{-\infty}^{-E_1} g(E) dE \approx 0,2 \frac{m^{*3/2}}{\hbar^3} \left(\frac{e^2 n_c^{1/3}}{\varepsilon} \right)^{3/2} V.$$

Это полное число связанных состояний возрастает как $n_c^{1/2}$, а не как n_c , поскольку область энергий, в которой электронные состояния различных центров перекрывается, увеличивается с ростом n_c .

Таким образом, плотность состояний в кристаллах с примесями имеет хвост в область запрещенных энергий.

§10 Фотоионизация дефектов

Электронные состояния дефектов вносят вклад в мнимую часть диэлектрической проницаемости. В недеформируемой решетке дефекты должны приводить к дельтаобразным максимумам в $\varepsilon_2(\omega)$. Однако такого рода переходы практически не наблюдаются в реальных кристаллах из-за электрон-фонного взаимодействия, приводящего к уширению и смещению линии поглощения. Поэтому детально поглощение на дефектах будет рассмотрено в следующей главе. Здесь мы рассмотрим переходы с дефектных уровней в состояния кристалла, которые приводят к широким полосам поглощения. Такого рода переходы возможны в случае фотоионизации дефектов, когда электроны с локализованного электронного состояния дефекта с известной энергией переходят в делокализованное состояние в зоне проводимости, или пустой уровень заполняется электроном с образованием делокализованного состояния в валентной зоне. В этом случае поглощение пропорционально концентрации дефектов. Мнимая часть диэлектрической проницаемости имеет вид

$$\varepsilon_2(\omega) = \frac{4\pi^2 e^2 N}{\hbar} \sum_{if} |\langle f | \mathbf{r} e | i \rangle|^2 \delta(\omega - \omega_{fi}). \quad (10.1)$$

Для оценки вклада дефекта в поглощение в качестве начального состояния можно рассматривать волновую функцию основного состояния дефекта, в качестве конечного — блоховскую функцию, принадлежащую континууму зоны проводимости $|c\mathbf{k}\rangle$. При этом $\omega_{fi} = (E_c(\mathbf{k}) - E_i)/\hbar$, где $E_c(\mathbf{k})$ — энергия зоны проводимости в зависимости от волнового вектора \mathbf{k} , а E_i — энергия основного состояния дефекта.

Для оценки матричного элемента дипольного перехода, входящего в (10.1), необходимо знать волновую функцию начального состояния. В общем случае эта волновая функция

неизвестна. В общем случае эта волновая функция неизвестна. Она легко может быть найдена в некоторых предельных случаях, например, в приближении эффективной массы, рассматривавшемся в разделе 6.2. Как уже указывалось, это приближение справедливо только для мелких уровней. Однако для вычисления матричного элемента $\langle f | \mathbf{re} | i \rangle$ достаточно знать асимптотику волновой функции $|i\rangle$ на достаточно больших расстояниях от дефекта, поскольку область малых расстояний относительно слабо влияет на величину матричного элемента перехода из-за делокализованного характера волновой функции конечного состояния. Для вычисления асимптотики $|i\rangle$ можно применить так называемый метод квантового дефекта [88–90].

На больших расстояниях от дефекта кулоновский потенциал остова, действующий на электрон, слабо меняется с расстоянием. В этом случае применимо приближение эффективной массы, и для волновой функции электрона на больших расстояниях от центра получаем уравнение

$$\left[E_i(0) - \frac{\hbar^2 \nabla^2}{2m_i^*} - \frac{(Z+1)e^2}{\epsilon r} - E_i \right] \psi_i(\mathbf{r}) = 0. \quad (10.2)$$

Предполагается, что состояние дефекта формируется из невырожденной изотропной зоны i с эффективной массой m_i^* . В этом уравнении Ze — заряд дефекта в неионизованном состоянии, E_i — точная энергия электронного состояния дефекта. Поскольку уравнение (10.2) справедливо только на больших расстояниях от дефекта, а на малых расстояниях потенциал может существенно отличаться от кулоновского, из этого уравнения невозможно определить энергию состояния дефекта E_i . Наоборот, зная точную энергию E_i , можно определить асимптотику волновой функции.

Основное состояние дефекта соответствует s -состоянию, и поэтому угловая часть сводится просто к $(4\pi)^{-1/2}$. Регулярное на бесконечности решение уравнения (10.2) имеет вид

$$\psi_i(\mathbf{r}) = \frac{2^\nu}{(4\pi)^{1/2} (\nu a^*)^{\nu+1/2} \Gamma(\nu+1)} r^{\nu-1} \exp(-r/\nu a^*), \quad (10.3)$$

где $a^* = \hbar^2 \epsilon / e^2 m_i^* (Z+1)$ — эффективный борковский радиус, а эффективное квантовое число ν определяется из

соотношения между эффективной постоянной Ридберга $Ry^* = (Z + 1)e^2/2\epsilon a^*$ и точной энергией E_i : $E_i = -Ry^*/\nu^2$.

Поскольку уравнение (10.2) неприменимо при малых r , волновая функция «плохо» ведет себя при $r \rightarrow 0$. Нормировка огибающей волновой функции основного состояния учитывает только вклад от больших расстояний. Естественно, что этот метод неприменим для однократно отрицательно заряженных дефектов, поскольку при отрыве электрона получается нейтральный остов без дальнедействующего потенциала.

В пределе чисто кулоновского взаимодействия, когда короткодействующая часть потенциала отсутствует, $E_i = -Ry^*$ и, следовательно, $\nu = 1$. Огибающая волновой функции на больших расстояниях меняется как e^{-r/a^*} . В случае короткодействующего притягивающего потенциала образуется глубокий уровень, энергия которого $|E_i|$ много больше Ry^* . Следовательно, $\nu \rightarrow 0$ и огибающая имеет вид

$$r^{-1} \exp(-r/a^* \nu) = r^{-1} \exp\left[-r\sqrt{2m_i^* |E_i|/\hbar}\right]$$

и, естественно, не зависит ни от заряда дефекта, ни от диэлектрической проницаемости. Эта волновая функция совпадает с решением уравнения Шредингера для дельтаобразной ямы. Таким образом, рассматриваемый метод дает возможность интерполировать результаты для разного вклада дальнедействующего кулоновского потенциала в полный потенциал взаимодействия в случае изменения параметра ν от 0 до 1.

Зона i , формирующая состояние дефекта, в принципе может не совпадать с зоной проводимости c , которой соответствуют делокализованные состояния. Введем энергию ионизации E_I , отсчитываемую от дна зоны проводимости: $E_I = E_c(0) - E_i$. Будем предполагать, что экстремум зоны i и зоны проводимости c находятся в Γ -точке. Волновая функция дефекта имеет вид $|i\rangle = \psi_i(\mathbf{r})|0, i\rangle$, где $|0, i\rangle$ — функция Блоха i -й зоны с равным нулю волновым вектором. Матричный элемент дипольного перехода можно преобразовать следующим образом:

$$\begin{aligned}
e\langle f | \mathbf{r} | i \rangle &= e \int d^3 r u_{c\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{r}} \mathbf{r} \psi_i(\mathbf{r}) u_{i\mathbf{0}}(\mathbf{r}) = \\
&= e \sum_{\mathbf{l}} e^{-i\mathbf{k}\mathbf{l}} \psi_i(\mathbf{l}) \int_v d^3 r u_{c\mathbf{k}}(\mathbf{r}) \mathbf{r} u_{i\mathbf{0}}(\mathbf{r}) = \sum_{\mathbf{l}} e^{-i\mathbf{k}\mathbf{l}} \psi_i(\mathbf{l}) \mathbf{d}_{ci}(\mathbf{0}).
\end{aligned}$$

Здесь была учтена плавность функции $\psi_i(\mathbf{r})$ и интегрирование ограничено по элементарной ячейке v . Матричный элемент дипольного момента $\mathbf{d}_{ci}(\mathbf{k})$ определен в параграфе 2. Переходя от суммирования по \mathbf{l} к интегрированию по непрерывной переменной \mathbf{r} , получим

$$e\langle f | \mathbf{r} | i \rangle = \mathbf{d}_{ci}(\mathbf{0}) \int d^3 r e^{-i\mathbf{k}\mathbf{r}} \psi_i(\mathbf{r}),$$

то есть матричный элемент фотоионизации дефекта определяется фурье-образом огибающей функции $\psi_i(\mathbf{r})$. Используя выражение (10.3), можно получить

$$e\langle f | \mathbf{r} | i \rangle = \mathbf{d}_{ci}(\mathbf{0}) \frac{2^v (v a^*)^{3/2} \sin[(v+1) \arctg(vka^*)]}{vka^* [1 + (vka^*)^2]^{(v+1)/2}},$$

где введено абсолютное значение импульса электрона в зоне проводимости $k = \sqrt{2m_c^*(\hbar\omega - E_I)}$. Удобно ввести безразмерную величину $X = (vka^*)^2 = 2m_c^*(\hbar\omega - E_I)/m_i^* E_i$. В случае водородоподобного дефекта ($v = 1$) вклад в $\varepsilon_2(\omega)$ пропорционален $\sqrt{X}/(1+X)^4$, в случае короткодействующего потенциала со слабым кулоновским вкладом $-v^2 \sqrt{X}/(1+X)^2$. Как видим, на пороге фотоионизации при $X \ll 1$ вклад в ε_2 обладает особенностью Ван-Хова, которая, естественно, связана с особенностью плотности состояний зоны проводимости. Поглощение растет до значений порядка $X \sim 1$ (то есть $\hbar\omega - E_I \sim m_i^* E_i/m_c^*$). В дальнейшем наступает спад вклада в ε_2 , более резкий в водородоподобном случае мелкого дефекта и более слабый в случае короткодействующего потенциала (рис. 28).

При этом рассмотрении не учитывались эффекты электронно-дырочных корреляций, приводящих в случае межзонных переходов к экситонным эффектам. В действительности волновая функция конечного состояния $|f\rangle$ не является плоской волной. Можно и для конечного состояния ввести огибающую функцию, которая в рамках приближения эффективной

массы имеет водородоподобный характер для непрерывного спектра. Вычисление матричных элементов в этом случае приводит к интегралу, включающему не только $\psi_i(\mathbf{r})$, но и водородоподобную функцию непрерывного спектра, соответствующую волновому вектору \mathbf{k} . Аналогичные интегралы возникают при исследовании фотоионизации атома водорода. В этой задаче [91] показано, что на пороге фотоионизации зависимость вклада в ε_2 пропорциональна не \sqrt{X} , а величине

$$\frac{\exp(-4\sqrt{2/X} \operatorname{arctg}(\sqrt{2/X}))}{1 - \exp(-2\pi\sqrt{2/X})}.$$

Эта величина имеет конечное значение при $X \rightarrow 0$. Конечное значение $\varepsilon_2(\omega)$ на порге фотоионизации не означает скачка поглощения. В действительности для заряженных дефектов без экранирования, к которым только и применимо все вышеизложенное, существует бесконечное число уровней ниже порога ионизации. Интенсивность перехода в эти квазидискретные состояния такова, что, как и для экситона Ванье – Мотта, на пороге ионизации никаких особенностей в $\varepsilon_2(\omega)$ не возникает.

Как уже отмечалось, вышеизложенное не совсем применимо к фотоионизации центров типа X^- , поскольку отрываемый электрон связан с остовом лишь короткодействующим потенциалом. Наличие такой потенциальной ямы может приводить к усилению поглощения сразу за порогом и к смещению максимума поглощения в область меньших энергий по сравнению со случаем $v \rightarrow 0$ предыдущего рассмотрения. Поскольку число уровней в короткодействующем потенциале конечно, особенности Ван-Хова при поглощении с фотоионизацией таких центров должны проявляться.

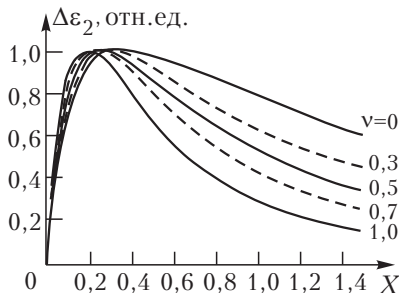


Рис. 28. Зависимость сечения фотоионизации от энергии над порогом для разной доли дальнедействующих сил: от кулоновского $v = 1$ до короткодействующего $v \rightarrow 0$ потенциала.

§11 Неупорядоченные системы

Дефекты кристаллической решетки имеются во всех кристаллах. Часто их концентрация становится настолько большой, что становится бессмысленно говорить о трансляционной симметрии. Увеличение концентрации дефектов в первую очередь нарушает дальний порядок. Однако для умеренных концентраций ближний порядок меняется слабо (например, не меняется координационное число для атомов кристалла). Для большой концентрации дефектов нарушается и ближний порядок, и твердое тело становится стеклом или аморфным веществом.

Теория и приложения неупорядоченных систем являются современной и быстро развивающейся областью науки, и им посвящено большое число книг ([6, 63, 92, 93]). Здесь мы обсудим только некоторые аспекты поведения этих интересных систем, которые важны в спектроскопии твердых тел.

Первым очевидным различием между системой без дальнего порядка и кристаллом является невозможность введения квазиимпульса, поскольку нарушается трансляционная симметрия, и поэтому теорема Блоха становится несправедливой. Поэтому невозможно ввести понятие зоны Бриллюэна, закона дисперсии и все другие понятия, являющиеся основой спектроскопии кристаллов. Однако имеется большой класс неупорядоченных систем, для которых эти понятия в некоторой степени применимы. Это системы с ближним порядком, для которых можно ввести характеристические размеры d квазикристаллических областей (для систем с точечными дефектами, дислокациями, двумерными дефектами и проч.) Эти размеры зависят от концентрации дефектов, например, для точечных дефектов с концентрацией n_c этот размер порядка $n_c^{-1/3}$. Тем самым в областях с линейными размерами d можно ввести квазиволновой вектор \mathbf{k} , неопределенность задания которого $\Delta\mathbf{k}$ зависит от d :

$$|\Delta \mathbf{k}| \approx \pi/d. \quad (11.1)$$

Если среднее межатомное расстояние a много меньше d , то эта неопределенность много меньше максимального волнового вектора (который порядка π/a). Для описания большого числа явлений в таких слабо неупорядоченных системах можно использовать картину, основанную на свойствах идеального кристалла. Состояния кристалла могут являться начальным приближением. Такие состояния рассеиваются дефектами и тем самым приобретают конечное время жизни. Неопределенность энергии ΔE порядка

$$\Delta E \approx \left| \frac{\partial E}{\partial \mathbf{k}} \Delta \mathbf{k} \right| \approx \hbar |\mathbf{v}| \frac{\pi}{d} \approx \frac{\pi \hbar}{\Delta t}, \quad (11.2)$$

где использовано определение скорости возбуждения $\mathbf{v} = \hbar^{-1} \partial E / \partial \mathbf{k}$ и соотношение (11.1). Интервал времени Δt соответствует времени, необходимому возбуждению на пересечение области размером порядка d , $\Delta t = d/|\mathbf{v}|$. Соотношение (11.2) может быть также получено из соотношения неопределенностей между энергией и временем. Этот результат показывает, что все сингулярности в спектре энергий (в частности, сингулярности Ван Хова) сглаживаются путем усреднения по интервалу энергии порядка (11.2).

Спектры оптического поглощения таких систем в области фундаментального поглощения близки к спектрам кристаллов, за исключением отсутствия выраженных особенностей.

Если степень неупорядоченности достаточно высока, то в системе может не наблюдаться и ближний порядок. Для таких систем практически невозможно использовать понятие волнового вектора и других квантовых чисел, традиционно используемых в физике кристаллов. Единственной функцией, которая может характеризовать спектр энергий, является плотность состояний $g(E)$. Как выше отмечалось, для слабо неоднородных систем плотность состояний незначительно отличается от плотности состояний для кристалла. Наибольшие отличия проявляются в зонах запрещенных энергий и вблизи от краев зон. Примером этого служит рассмотренная в разделе 9.4 плотность состояний для заряженных дефектов. Хвосты плотности состояний появляются у каждого края зон. Этот пример является одним из возможных изменений плотности состояний. Если дефекты глубокие, нейтральные и т.д., то

результаты отличаются от описываемых формулой (9.4). Хвосты плотности состояний могут иметь другую зависимость от энергии (например, экспоненциальную), они могут быть немонотонными и проч. Тем самым в зоне, бывшей запрещенной, могут наблюдаться новые пики плотности состояний. Для сильно неупорядоченных систем хвосты от дна зоны проводимости и от потолка валентной зоны могут перекрываться. Тем самым никакой запрещенной зоны (в прежнем смысле этого понятия) может не наблюдаться.

Несмотря на это, даже системы с перекрывающимися хвостам плотности состояний могут быть диэлектриками. Это связано с тем, что непрерывность плотности состояний не гарантирует делокализованный характер волновых функций. Глубокие состояния оказываются пространственно-разделенными, и перенос возбуждения с одного состояния на другое имеет малую вероятность. Определение границы подвижности возбуждений является предметом рассмотрения теории перколяции (протекания).

Спектр поглощения неупорядоченной системы может быть рассчитан аналогично тому, как это делалось для кристаллов. Формулы (3.8) и (3.32) могут быть выведены без каких-либо предположений о трансляционной инвариантности системы. Единственным предположением является макроскопическая пространственная однородность системы. Следствием теоремы Блоха для кристалла является то, что входящий в (3.32) квадрат матричного элемента $\left(\left| \mathbf{D}_{nm} \right|^2 \right)_{-\mathbf{k}}$ отличен от нуля только для состояний, удовлетворяющих соотношению $\mathbf{k}_n = \mathbf{k}_m + \mathbf{k}$. Для неупорядоченных систем это несправедливо. В соответствии с вышесказанным, для слабо неупорядоченных систем волновой вектор удовлетворяет соотношению

$$|\mathbf{k}_n - \mathbf{k}_m - \mathbf{k}| \leq \pi/d ,$$

в то время как для сильно неупорядоченных систем на волновые вектора не накладывается никаких ограничений. В последнем предельном случае естественно заменить матричный элемент константой. Это приводит к известному соотношению

$$\varepsilon_2(\omega) \sim \int dE g(E) f(E) g(E + \hbar\omega) [1 - f(E + \hbar\omega)] , \quad (11.3)$$

то есть диэлектрическая проницаемость пропорциональна

свертке плотностей состояний для свободной $g(E + \hbar\omega)$ и заполненной $g(E)$ зон. Поскольку $g(E)$ для делокализованных состояний близки к плотностям состояний кристалла, можно использовать приближение $g(E) \sim \sqrt{E - E_g}$ для зоны проводимости и $g(E) \sim \sqrt{-E}$ для валентной зоны вблизи их дна и потолка, соответственно. Отсюда следует, что $\varepsilon_2(\omega) \sim (\hbar\omega - E_g)^2$ вблизи порога фундаментального поглощения ($\hbar\omega > E_g$). Этот результат показывает, что особенности Ван Хофа в спектре поглощения пропадают.

Ниже области фундаментального поглощения формула (11.3) переоценивает мнимую часть диэлектрической проницаемости. Это связано с тем, что различные локализованные состояния пространственно могут не перекрываться, и тем самым матричный элемент отличен от нуля только для состояний с перекрывающимися волновыми функциями. Тем самым в (11.3) проявляется корреляция заполненных и свободных состояний.

Рассмотрим два различных случая систем с параметрами, которые медленно меняются в пространстве. Для таких систем каждый физически малый объем можно рассматривать как идеальный кристалл, но параметры этих кристалликов меняются от одной точки к другой. Первый случай соответствует пространственному изменению запрещенной зоны. Дно зоны проводимости $U_c(\mathbf{r})$ и потолок валентной зоны $U_v(\mathbf{r})$ могут плавно изменяться. Тем самым запрещенная зона $E_g(\mathbf{r}) = U_c(\mathbf{r}) - U_v(\mathbf{r})$ также меняется от точки к точке. Поэтому диэлектрическая проницаемость такого кристалла может быть записана в форме

$$\varepsilon_2(\omega) = \int d\Delta P(\Delta) \varepsilon_2^0(\omega; E_g + \Delta), \quad (11.4)$$

где $\varepsilon_2^0(\omega; E_g)$ — диэлектрическая проницаемость кристалла с запрещенной зоной E_g , а $P(\Delta)$ — вероятность того, что разница $E_g(\mathbf{r}) - E_g$ между микроскопической и средней запрещенной зонами равна Δ . Например, в случае плавных флуктуаций с гауссовским распределением хвост $\varepsilon_2(\omega)$ также имеет гауссовский характер: $\varepsilon_2(\omega) \sim \exp[-C(E_g - \hbar\omega)^2]$, где C — константа, уменьшающаяся с увеличением степени неупорядоченности.

Такие некоррелированные изменения $U_c(\mathbf{r})$ и $U_v(\mathbf{r})$ могут

быть вызваны короткодействующим флуктуационным потенциалом, связанным, например, с изменениями постоянной решетки. Если сдвиг зон связан с флуктуациями заряженных областей в системе, электрическое поле действует на электроны в зоне проводимости и в валентной зоне одинаково. Поэтому в этом случае $U_c(\mathbf{r}) = U_v(\mathbf{r}) + E_g$, и отсюда $E_g(\mathbf{r}) = E_g$. Тем самым формула (11.4) становится неприменимой. Другой подход к этой проблеме может привести в получению результата. Заряженные области в кристалле создают электрическое поле $\mathcal{E}(\mathbf{r})$. Поэтому необходимо применять выражения для диэлектрической проницаемости в электрическом поле (8.5). Тем самым диэлектрическая проницаемость системы с заряженными областями может быть записана в виде

$$\varepsilon_2(\omega) = \int d\mathcal{E} P(\mathcal{E}) \varepsilon_2^0(\omega; \mathcal{E}), \quad (11.5)$$

где $P(\mathcal{E})$ — вероятность обнаружить электрическое поле \mathcal{E} в образце. В случае случайно распределенных заряженных примесей с концентрацией n_c распределение $P(\mathcal{E})$ является распределением Хольцмарка (см., например, [48]):

$$P(\mathcal{E}) = \mathcal{E}_0^{-1} H(\mathcal{E}/\mathcal{E}_0),$$

где $\mathcal{E}_0 = 2,6 en_c^{2/3}$, то есть имеет порядок поля на расстоянии $n_c^{-1/3}$ от примеси, а нормированная функция $H(\beta)$ равна

$$H(\beta) = 2\pi^{-1} \beta \int_0^{\infty} dx x \sin(\beta x) e^{-x^{-3/2}}$$

и имеет максимум вблизи $\beta \approx 1,6$, и асимптотическое поведение $H(\beta) \sim \beta^2$ при $\beta \ll 1$ и $H(\beta) \sim \beta^{-5/2}$ при $\beta \gg 1$.

Для плавных полей с гауссовым распределением диэлектрической проницаемости имеет вид $\varepsilon_2(\omega) \sim \exp[-G(E_g - \hbar\omega)]$, то есть ведет себя экспоненциально в области энергий $\hbar\omega < E_g$. G в этой формуле является постоянной [6]. Такая зависимость диэлектрической проницаемости ниже края фундаментального поглощения известна как экспериментально обнаруженное правило Урбаха. Этот закон был обнаружен для кристаллов [94], для которых он имеет другое объяснение (см. третью главу). Однако аналогичная экспоненциальная зависимость позволяет использовать то же наименование и для неупорядоченных систем.

Глава 3

Электронные состояния в деформируемой решетке

§12 Фононы в кристаллах

12.1 Фононы в идеальном кристалле

Вернемся к рассмотрению исходного гамильтониана кристалла (4.1). В предыдущей главе пренебрегалось движением ядерной подсистемы. Здесь мы учтем влияние движения ядер на электронные свойства. Запишем этот гамильтониан в виде

$$\hat{H}_0 = \hat{T}_e + \hat{T}_N + \hat{V}(\mathbf{r}, \mathcal{R}), \quad (12.1)$$

где \mathbf{r} и \mathcal{R} — совокупность координат всех электронов и ядер, \hat{T}_N и \hat{T}_e — кинетические энергии ядер и электронов, а $\hat{V}(\mathbf{r}, \mathcal{R})$ — оператор потенциальной энергии. Исследуем более подробно приближение по малому параметру m/M , где m и M — массы электрона и ядра. Частоты движения электронов намного выше, чем ядер. Поэтому электроны в каждый момент находятся как бы в статическом поле ядерной подсистемы. Напротив, ядра «чувствуют» только усредненный по быстрому движению электронов потенциал. Наиболее последовательно это отражено в идее адиабатического приближения, или приближения Борна — Оппенгеймера. В этом приближении волновая функция ищется в невырожденном случае в виде произведения функций, зависящих от электронных и

ионных координат:

$$\Phi(\mathbf{r}, \mathcal{R}) = \psi(\mathbf{r}; \mathcal{R}) \chi(\mathcal{R}) \quad (12.2)$$

Электронная волновая функция $\psi(\mathbf{r}; \mathcal{R})$ зависит от ионных координат \mathcal{R} как от параметров. Фиксируя координаты ядерной подсистемы \mathcal{R} и полагая $\hat{\mathcal{T}}_N = 0$, получают уравнение для $\psi(\mathbf{r}; \mathcal{R})$:

$$[\hat{\mathcal{T}}_e + \hat{\mathcal{V}}(\mathbf{r}, \mathcal{R})] \psi(\mathbf{r}; \mathcal{R}) = E_e(\mathcal{R}) \psi(\mathbf{r}; \mathcal{R}) . \quad (12.3)$$

Здесь $E_e(\mathcal{R})$ — собственное значение уравнения (12.3) для данной конфигурации ядер \mathcal{R} . В предыдущей главе это уравнение исследовалось для равновесной конфигурации \mathcal{R}_0 . Уравнение (12.3) сложнее последнего хотя бы тем, что при $\mathcal{R} \neq \mathcal{R}_0$ частично или полностью снимается трансляционная и точечная симметрия системы.

Уравнение для функции $\chi(\mathcal{R})$ можно получить следующим способом. Применим оператор $\hat{\mathcal{H}}_0$ (12.1) к функции (12.2):

$$\begin{aligned} \hat{\mathcal{H}}_0 \Phi(\mathbf{r}, \mathcal{R}) &= E_e(\mathcal{R}) \psi(\mathbf{r}; \mathcal{R}) \chi(\mathcal{R}) + \sum_i \frac{\hat{\mathbf{P}}_i^2}{2M_i} \psi(\mathbf{r}; \mathcal{R}) \chi(\mathcal{R}) = \\ &= \psi(\mathbf{r}; \mathcal{R}) \left[E_e(\mathcal{R}) + \sum_i \frac{\hat{\mathbf{P}}_i^2}{2M_i} \right] \chi(\mathcal{R}) + \\ &+ \left\{ \chi(\mathcal{R}) \sum_i \frac{\hat{\mathbf{P}}_i^2}{2M_i} \psi(\mathbf{r}; \mathcal{R}) + \sum_i \frac{1}{M_i} [\hat{\mathbf{P}}_i \chi(\mathcal{R})][\hat{\mathbf{P}}_i \psi(\mathbf{r}; \mathcal{R})] \right\}. \end{aligned} \quad (12.4)$$

В теории адиабатического приближения предполагается, что члены в фигурной скобке малы, поскольку $\psi(\mathbf{r}; \mathcal{R})$ зависит от \mathcal{R} относительно слабо. Тогда получаем уравнение для определения полной энергии и волновой функции решетки:

$$\left[\sum_i \frac{\hat{\mathbf{P}}_i^2}{2M_i} + E_e(\mathcal{R}) \right] \chi(\mathcal{R}) = E \chi(\mathcal{R}) .$$

Это же уравнение можно вывести и из вариационного принципа, используя функцию (12.2) в качестве пробной функции. В энергию $E_e(\mathcal{R})$ входят как энергия электростатического

взаимодействия ионных остовов (член $\sum_{i \neq j} \frac{Z_i Z_j e^2}{|\mathbf{R}_i - \mathbf{R}_j|}$), так и

энергия взаимодействия между остовом и электронными оболочками.

Как выше уже указывалось, \mathcal{R} представляет собой совокупность координат всех атомов. Если в основной области кристалла находится N элементарных ячеек, а каждая ячейка состоит из σ атомов, то \mathcal{R} имеет $3N\sigma$ компонент. Введем $N\sigma$ трехмерных векторов отклонений $\mathbf{u}_{\alpha\mathbf{n}}$ (α – номер атома в элементарной ячейке, соответствующей радиус-вектору \mathbf{n}). Тогда $E_e(\mathcal{R})$ можно разложить в ряд по отклонениям:

$$E_e(\mathcal{R}) = E_e(\mathcal{R}) + \sum_{\mathbf{n}\alpha} \mathbf{B}_{\mathbf{n}\alpha} \mathbf{u}_{\mathbf{n}\alpha} + \sum_{\mathbf{n}\alpha\mathbf{n}'\alpha'} A_{i\mathbf{n}\alpha; j\mathbf{n}'\alpha'} u_{i\mathbf{n}\alpha} u_{j\mathbf{n}'\alpha'} + \dots$$

Высшие члены, ответственные за аангармоничность колебаний решетки, можно опустить вследствие их малости в большинстве интересных для спектроскопии случаев. В идеальном кристалле матрица $A_{i\mathbf{n}\alpha; j\mathbf{n}'\alpha'}$ может зависеть лишь от $\mathbf{n} - \mathbf{n}'$. Вследствие этого удобно перейти к Фурье-компонентам смещений:

$$\bar{\xi}_{\alpha}(\mathbf{k}) = N^{-1/2} \sum_{\mathbf{n}} e^{-i\mathbf{k}\mathbf{n}} \mathbf{u}_{\mathbf{n}\alpha},$$

где \mathbf{k} принимает N дискретных значений из первой зоны Бриллюэна. Переходя и в волновой функции к новым координатам, можно получить оператор Гамильтона, действующий на χ , в виде матрицы $3\sigma \times 3\sigma$, компоненты которой зависят от волнового вектора \mathbf{k} . Диагонализация этой матрицы приведет к 3σ уравнениям для каждого \mathbf{k} , определяющих нормальные колебания, называемые модами, с собственными значениями $\Omega_s^2(\mathbf{k})$. Каждое из этих уравнений совпадает с квантовым уравнением для осциллятора с частотой $\Omega_s(\mathbf{k})$, где s – номер ветви колебаний. Вектор Фурье-компонент смещений \mathbf{e}_{α}^s , соответствующий собственному значению $\Omega_s(\mathbf{k})$, определяет смещение α -атома решетки при колебаниях моды s . Смещение какого-либо атома будет выражаться через нормальные координаты $Q_{\mathbf{k}s}$ этого осциллятора с помощью преобразования

$$\mathbf{u}_{\mathbf{n}\alpha} = N^{-1/2} \sum_{\mathbf{k}s} \mathbf{e}_{\alpha}^s(\mathbf{k}) Q_{\mathbf{k}s} e^{i\mathbf{k}\mathbf{n}}. \quad (12.5)$$

В кристаллах, содержащих один атом в элементарной ячейке, существуют три ветви колебаний. Для всех этих ветвей $\lim_{\mathbf{k} \rightarrow \mathbf{0}} \Omega_s(\mathbf{k}) = 0$. Обладающие таким свойством колебания кристалла называются акустическими, поскольку они связаны с распространением звука в твердом теле. В кубических кристаллах колебания можно разделить на продольные, связанные с изменением объема элементарной ячейки (одна ветвь), и поперечные (две ветви). Эти ветви называются LA - и TA -ветвями колебаний. Если элементарная ячейка содержит несколько атомов, то появляются, в дополнение к трем акустическим ветвям, еще $3(\sigma - 1)$ ветвей. Для них уже $\Omega_s(\mathbf{0}) \neq 0$. Поскольку частоты $\Omega_s(\mathbf{k})$ этих ветвей лежат в инфракрасном диапазоне, такие колебания называются оптическими. В кубических кристаллах их также можно подразделить на продольные LO - и поперечные TO -колебания.

Широко распространены следующие аппроксимации частот колебаний. Для акустических ветвей в значительной области зоны Бриллюэна справедливо соотношение $\Omega_s(\mathbf{k}) = c_s |\mathbf{k}|$, где c_s — скорость распространения продольных (для LA -моды) или поперечных (для TA -мод) звуковых волн. Для оптических мод, которые в пределе $\mathbf{k} \rightarrow \mathbf{0}$ приводят лишь к смещениям атомов в элементарной ячейке по отношению друг к другу, но не к изменениям объема ячейки, часто удобно применять аппроксимацию Эйнштейна $\Omega_s(\mathbf{k}) = \Omega_s(\mathbf{0})$. В ионных кристаллах необходимо учитывать возникновение продольного электрического поля поляризации при LO -колебаниях. Это явление аналогично LT -расщеплению экситонов (см. раздел 6.2) и вызывает расщепление частот LO - и TO -колебаний, причем частоты продольных Ω_{LO} и поперечных Ω_{TO} колебаний связаны между собой соотношением Лиддана-Сакса-Теллера

$$\Omega_{LO}^2 = \Omega_{TO}^2 \frac{\epsilon_0}{\epsilon_\infty}.$$

Здесь введены две диэлектрические проницаемости, часто встречающиеся при описании ионных кристаллов: статическая диэлектрическая проницаемость ϵ_0 , измеренная в постоянном поле, и высокочастотная диэлектрическая проницаемость ϵ_∞ , измеренная при такой частоте, при которой все электроны успевают следовать изменениям высокочастотного поля, а ионы не успевают, т. е. на частоте, много большей

частот фононного спектра, но много меньшей частоты электронных переходов (то есть в области прозрачности диэлектрического кристалла). В ϵ_∞ , в отличие от ϵ_0 , не вносит вклад поляризуемость ионной подсистемы. Более подробно это различие проанализировано в разделе 14.3.

По аналогии с плотностью электронных состояний можно ввести плотность фононных состояний ρ , которая является функцией частоты Ω :

$$\rho(\Omega) = \sum_{\mathbf{k}s} \delta(\Omega - \Omega_s(\mathbf{k})) .$$

Эта функция $\sum_{\mathbf{k}s}$ должна иметь не менее четырех особенностей Ван Хова для каждой ветви (причины такого поведения полностью аналогичны приведенным для плотности электронных состояний $g(E)$ в разделе 4.5). Все сингулярности имеют ту же функциональную зависимость, что и для $g(E)$, то есть $\sqrt{|\Omega - \Omega_{\text{ext}}|}$, за исключением точки $\Omega = 0$. В этой точке из-за линейной зависимости $\Omega_s(\mathbf{k}) = c_s |\mathbf{k}|$ для акустических мод $\rho(\Omega) \sim \Omega^2$ (рис. 29).

Фононные состояния могут быть вырождены в точках высокой симметрии зоны Бриллюэна. В Γ -точке ($\mathbf{k} = 0$) акустические моды трехкратно вырождены. Степень вырождения других мод зависит от симметрии кристалла. Как и для электронных состояний, имеются невырожденные, двукратно и трехкратно вырожденные фононные моды.

После введения нормальных колебаний задача классификации состояний фононного спектра и описания его свойств, а также свойств волновых функций существенно упрощается. Гамильтониан, действующий на координаты ионных остовов, распадается в нормальных координатах на сумму операторов Гамильтона для осцилляторов:

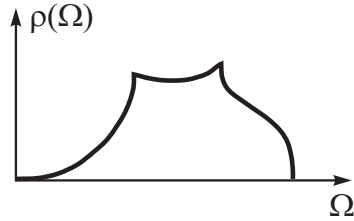


Рис. 29. Плотность состояний для ветви акустических фононов

$$\begin{aligned}
\hat{\mathcal{H}}_{\text{ph}} &\equiv \sum_i \hat{\mathbf{P}}_i^2 / 2M_i + E_e(\mathcal{R}) = \\
&= \frac{1}{2} \sum_{\mathbf{k}s} [\hat{P}_{\mathbf{k}s} \hat{P}_{-\mathbf{k}s} + B_{\mathbf{k}s} \hat{Q}_{\mathbf{k}s} + \Omega_s^2(\mathbf{k}) \hat{Q}_{\mathbf{k}s} \hat{Q}_{-\mathbf{k}s}] + E_e(\mathcal{R}_0) = \\
&= \sum_{\mathbf{k}s} \hbar \Omega_s(\mathbf{k}) (b_{\mathbf{k}s}^+ b_{\mathbf{k}s} + \frac{1}{2}) + \\
&\quad + \sum_{\mathbf{k}s} B_{\mathbf{k}s} \sqrt{\frac{\hbar}{2\Omega_s(\mathbf{k})}} (b_{\mathbf{k}s}^+ + b_{-\mathbf{k}s}) + E_e(\mathcal{R}_0),
\end{aligned} \tag{12.6}$$

где $\hat{P}_{\mathbf{k}s}$ — импульсы, сопряженные нормальным координатам $\hat{Q}_{\mathbf{k}s}$, $b_{\mathbf{k}s}^+$, $b_{\mathbf{k}s}$ — операторы рождения и уничтожения фононов с импульсом \mathbf{k} , принадлежащих ветви s . Операторы $\hat{P}_{\mathbf{k}s}$ и $\hat{Q}_{\mathbf{k}s}$ связаны с $b_{\mathbf{k}s}^+$, $b_{\mathbf{k}s}$ соотношениями

$$\begin{aligned}
\hat{Q}_{\mathbf{k}s} &= \sqrt{\frac{\hbar}{2\Omega_s(\mathbf{k})}} (b_{\mathbf{k}s}^+ + b_{-\mathbf{k}s}), \\
\hat{P}_{\mathbf{k}s} &= i \sqrt{\frac{\hbar \Omega_s(\mathbf{k})}{2}} (b_{\mathbf{k}s}^+ - b_{-\mathbf{k}s}).
\end{aligned} \tag{12.7}$$

Волновая функция $\chi(\mathcal{R})$ распадается на произведение волновых функций каждой моды — решений одномерного уравнения Шредингера для осциллятора с частотой $\Omega_s(\mathbf{k})$. Собственная функция $\chi_n(Q)$ одномерного осциллятора характеризуется одним квантовым числом n , называемом числом заполнения соответствующей моды. Тем самым собственная функция $\chi(\mathcal{R})$ будет характеризоваться набором чисел заполнения $\{n_{\mathbf{k}s}\}$ для каждой моды $\mathbf{k}s$:

$$\chi(\mathcal{R}) = \prod_{\mathbf{k}s} \chi_{n_{\mathbf{k}s}}(Q_{\mathbf{k}s}). \tag{12.8}$$

Таким образом, собственная функция полного гамильтониана в адиабатическом приближении есть произведение (12.2) электронной волновой функции $\psi(\mathbf{r}; \{Q_{\mathbf{k}s}\})$ на волновую функцию $\chi(\{Q_{\mathbf{k}s}\})$. Поскольку $E_e(\mathcal{R})$ зависит от состояния электронной подсистемы, коэффициенты разложения энергии по смещениям $\mathbf{u}_{n\alpha}$ будут также зависеть от этого состояния. Это означает, что $B_{\mathbf{k}s}^e$ и $\Omega_s(\mathbf{k})$ могут отличаться для различных электронных состояний, и поэтому у

соответствующих коэффициентов будет указываться индекс состояния.

Обычно предполагают, что изменение электронного состояния не приводит к качественному изменению характера фоновго спектра. Часто используется более жесткое условие: предполагается, что сдвигается только положение равновесия, а частоты фоновных мод остаются неизменными. При этом во всем выражении для гамильтониана (12.6) при переходе от одного электронного состояния к другому происходит изменение только E_e и постоянных $B_{\mathbf{k}s}^e$. Такое приближение называется приближением линейной электрон-решеточной связи, поскольку изменение фоновго гамильтониана при переходе к другому электронному состоянию происходит только в линейных по $Q_{\mathbf{k}s}$ членах.

Силы электрон-фоновной связи для разных мод отличаются друг от друга. Если рассматривать только одну моду $\mathbf{k}s$, то можно построить диаграмму, изображенную на рис. 30. На ней параболы соответствуют двум различным электронным состояниям i и j . Минимумы $\bar{Q}_{\mathbf{k}s}^i$ и $\bar{Q}_{\mathbf{k}s}^j$ кривых потенциальной энергии находятся в точках $-B_{\mathbf{k}s}^i/2\Omega_s(\mathbf{k})$ и $-B_{\mathbf{k}s}^j/2\Omega_s(\mathbf{k})$. При переходе из одного электронного состояния в другое минимум потенциальной энергии смещается на величину $q_{ij,\mathbf{k}s} = \bar{Q}_{\mathbf{k}s}^j - \bar{Q}_{\mathbf{k}s}^i$. Как видно из выражений (12.7), эта величина размерна. Гораздо удобнее соответствующие изменения выражать через безразмерную величину. Для этого выразим изменение потенциальной энергии $\Delta E = q_{ij,\mathbf{k}s}^2 \Omega_s^2(\mathbf{k})/2$ при переходе из положения равновесия $\bar{Q}_{\mathbf{k}s}^i$ в $\bar{Q}_{\mathbf{k}s}^j$ без изменения электронного состояния в квантах $\hbar\Omega_s(\mathbf{k})$ соответствующей моды:

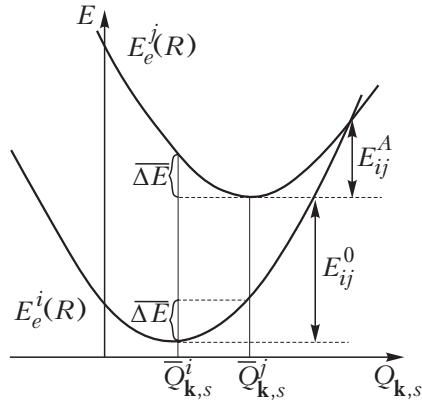


Рис. 30. Конфигурационная диаграмма для моды $\mathbf{k}s$ и соответствующие параметры.

$$S_{\mathbf{k}_s} = \overline{\Delta E} / \hbar \Omega_s(\mathbf{k}) = q_{ij, \mathbf{k}_s}^2 \Omega_s(\mathbf{k}) / 2\hbar. \quad (12.9)$$

Эта величина называется «фактором Хуанга—Риса». Как было показано независимо в работах [95, 96], величины (12.9) играют основную роль в формировании линии поглощения и испускания света.

12.2 Фононы в кристаллах с дефектами

Естественно, что наличие дефектов в кристалле изменяет не только электронные состояния, но и фононные. Дефекты снимают трансляционную симметрию кристалла, и поэтому волновой вектор \mathbf{k} перестает быть хорошим квантовым числом, описывающим фононы. В этом случае подробное описание фононов и их взаимодействия с электронной подсистемой сложнее, чем в идеальном кристалле, и поэтому мы отметим только основные отличия от предыдущего случая.

Если пренебречь электрон-фононным взаимодействием, то дефекты приводят к рассеянию фононов и к появлению локальных мод. Рассеяние фононов дефектом может быть описано в нулевом приближении, то есть в базисе фононов с определенным \mathbf{k} в идеальном кристалле. В таком приближении рассеяние вызывает переходы между различными модами (от \mathbf{k} -фононов к \mathbf{k}' -фононам). Этот эффект важен для описания термализации фононной подсистемы, термопроводимости и т.д. Для описания электрон-фононного взаимодействия рассеяние имеет меньшее значение.

Взаимодействие фононов с различными волновыми векторами \mathbf{k} может приводить к образованию локальных мод (можно сравнить с образованием локальных электронных состояний, см. уравнение (9.1)). Эти моды характеризуются частотами Ω_λ . Для локальных мод сдвиги атомов из положения равновесия уменьшаются с ростом расстояния от дефекта. В предельном случае можно рассматривать колебания самого дефекта и одной-двух координационных сфер вокруг него. Эти локальные моды могут классифицироваться с использованием точечной симметрии позиции дефекта. Поэтому могут существовать невырожденные состояния (A -моды), двукратно (E -моды) и трехкратно (T -моды) вырожденные состояния.

Каждая вырожденная мода описывается двумя (для E -мод) или тремя (в случае T -мод) нормальными координатами. Из-за вырождения эти координаты не определяются однозначно.

Взаимодействие электронных состояний с локальными фоновыми модами может быть описано аналогично описанию взаимодействия с делокализованными фоновыми модами. Основным различием между этими случаями для невырожденных электронных состояний является величина факторов Хуанга-Риса (12.9): в общем случае для локальных мод S_λ много больше, чем S_{k_s} (см. ниже, раздел 13.1).

Качественное отличие между этими случаями возникает в случае вырожденного электронного состояния дефекта, взаимодействующего с вырожденной фоновой модой. Ян и Теллер [97] показали, что для молекулы или комплекса в твердом теле с вырожденными электронными состояниями имеются такие искажения положений окружающих атомов, которые понижают энергию и симметрию системы. Эта теорема справедлива для большинства случаев, за исключением одномерных систем и крамерсова вырождения.

В результате эффекта Яна-Теллера уменьшается симметрия системы. В большинстве случаев это связано со смещением дефектного атома из дефектной межузельной позиции или позиции замещения.

Для вырожденных мод необходимо исследовать двумерную поверхность потенциальной энергии (для E -мод) или трехмерную гиперповерхность (для T -мод). Для простоты мы обсудим только случай двукратного вырождения. Локальная E -мода описывается двумя конфигурационными координатами Q_1 и Q_2 и частотой Ω_E . Потенциальная энергия двукратно вырожденного электронного E -состояния, взаимодействующего с E -модой, может быть записана в виде (см., например, [5])

$$\pm B_E^E \sqrt{Q_1^2 + Q_2^2} + \frac{1}{2} \Omega_E^2 (Q_1^2 + Q_2^2),$$

а для трехкратно вырожденного T -состояния —

$$2B_E^T \left\{ \begin{array}{c} (Q_1 + \sqrt{3}Q_2)/2 \\ (Q_1 - \sqrt{3}Q_2)/2 \\ -Q_1 \end{array} \right\} + \frac{1}{2} \Omega_E^2 (Q_1^2 + Q_2^2).$$

Эти потенциальные поверхности изображены на рис. 31 и 32.

В случае $E - E$ взаимодействия минимум потенциальной энергии достигается на окружности

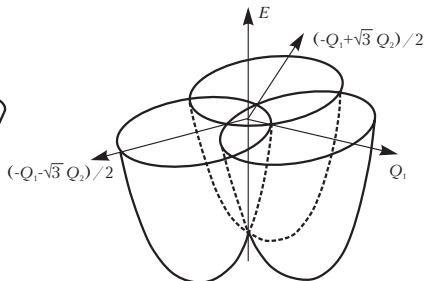
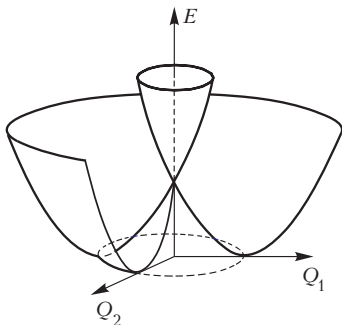


Рис. 31. Потенциальные поверхности в случае $E - E$ взаимодействия.

Рис. 32. Потенциальные поверхности в случае $T - E$ взаимодействия.

$$\rho = \sqrt{Q_1^2 + Q_2^2} = B_E^E / \Omega_E^2 ,$$

и энергия не зависит от «угла» $\varphi = \arctg(Q_1/Q_2)$. Двукратно вырожденное состояние при смещении из точки $Q_1 = Q_2 = 0$ расщепляется на два состояния. Понижение потенциальной энергии из-за эффекта Яна-Теллера для нижнего состояния равно

$$E_{JT} = -(B_E^E)^2 / 2\Omega_E^2 .$$

Потенциальные поверхности в случае $T - E$ взаимодействия представляют из себя три параболоида, сдвинутые из центра в точки

$$\rho = B_E^T / \Omega_E^2 , \quad \varphi = \{0, \pm 2\pi/3\} .$$

Энергия в этих минимумах равна

$$E_{JT} = -(B_E^T)^2 / 2\Omega_E^2 .$$

Поведение дефектного атома сильно различается в случаях взаимодействий $E - E$ и $T - E$. В случае $T - E$ дефект может быть найден в одной из трех эквивалентных позиций. Если кинетическая энергия меньше, чем E_{JT} , дефект оказывается локализованным в одной из потенциальных ям, и время от времени туннелирует из одной в другую (статический эффект Яна-Теллера). Если время туннелирования меньше

характерных времен эксперимента, или кинетическая энергия больше E_{JT} , дефект будет распределен между ямами (динамический эффект Яна–Теллера).

В случае взаимодействия $E - E$ вырождение минимума потенциальной энергии бесконечно (энергия не зависит от ϕ). Тем самым оказывается возможным «вращение» конфигурации дефекта с постоянным ρ , и движение дефекта и окружающих атомов не может быть сведено к их колебаниям. В этом случае динамический эффект Яна–Теллера играет определяющую роль.

Эффект Яна–Теллера важен не только в случае полностью вырожденного электронного состояния, но и в случае, если электронные состояния разнесены по энергии меньше, чем E_{JT} (иногда такой случай называется псевдоэффектом Яна–Теллера). Это важно учитывать, например, при наличии спин-орбитального расщепления электронных состояний.

§13 Переходы между состояниями дефектов при учете фононов

13.1 Функция формы линии поглощения при учете фононов

В предыдущем параграфе были введены все необходимые величины для описания взаимодействия света с электронной подсистемой в присутствии фононов. Для простоты вначале рассмотрим поглощение света с возбуждением локализованных электронных состояний. Такого рода состояния, если они не являются вырожденными, полностью характеризуются номером i в отличие от распределенных состояний, например экситонных, для которых помимо номера зоны необходимо указывать волновой вектор \mathbf{k} . Чтобы до предела упростить изложение, будем иметь в виду модель дефекта с двумя состояниями: основным i и возбужденным f . Мнимая часть диэлектрической проницаемости с точностью до поправки на эффективное поле при переходе из заполненного начального состояния в пустое конечное равна

$$\varepsilon_2(\omega) = 4\pi^2 e^2 N \hbar^{-1} \sum_{\{n_i\}, \{n_f\}} \left\langle \left\langle f \{n_f\} \left| \sum_j \mathbf{r}_j \mathbf{e} \right| i \{n_i\} \right\rangle \right\rangle^2 \times \quad (13.1)$$

$$\times f(E_{i\{n_i\}}) \delta(\omega - \omega_{f\{n_f\}, i\{n_i\}}).$$

Здесь состояние дефекта и окружающего его кристалла характеризуется индексом i или f и набором чисел заполнения $\{n_{i,f}\}$ фононных мод. Энергия кристалла с дефектом может быть записана в виде суммы энергии дефекта и суммы по всем модам энергий осцилляторов:

$$E_{i\{n_i\}} = E_i^0 + \sum_{\mathbf{k}s} (n_{\mathbf{k}s} + \frac{1}{2}) \hbar \Omega_s(\mathbf{k}),$$

где $n_{\mathbf{k}s}$ — числа из набора $\{n_i\}$. Поскольку фононы в кристалле без учета ангармонизмов являются идеальным бозе-газом, функция распределения может быть легко выражена через температуру T :

$$f(E_{i\{n_i\}}) = \prod_{\mathbf{k}s} \left[\exp\left(\frac{\hbar\Omega_s(\mathbf{k})n_{\mathbf{k}s}}{k_B T}\right) - 1 \right]^{-1}.$$

Здесь $k_B = 1,381 \cdot 10^{23}$ Дж/К = (1/11604) эВ/К — постоянная Больцмана. Из-за наличия функции распределения по начальным состояниям в (13.1) суммирование по ним сводится к усреднению по начальному состоянию с температурой T . Усреднение такого рода ниже будем обозначать скобками $\langle\langle \dots \rangle\rangle_T$.

Дальнейшее продвижение в вычислении ε_2 возможно только при упрощении матричного элемента в выражении (13.1). Пользуясь адиабатическим приближением, можно записать матричный элемент в виде

$$\begin{aligned} e \left\langle f\{n_f\} \left| \sum_j \mathbf{r}_j \mathbf{e} \right| i\{n_i\} \right\rangle &= \\ &= \int d\mathcal{R} \int dr \psi_f^*(r; \mathcal{R}) \chi_f^*(\mathcal{R}) \left(\sum_j \mathbf{r}_j \mathbf{e} \right) \psi_i(r; \mathcal{R}) \chi_i(\mathcal{R}) = \\ &= \int d\mathcal{R} \chi_f^*(\mathcal{R}) \mathbf{d}_{fi}(\mathcal{R}) \chi_i(\mathcal{R}), \end{aligned}$$

где через $\mathbf{d}_{fi}(\mathcal{R})$ обозначен дипольный момент перехода из начального состояния в конечное в заданной конфигурации \mathcal{R} . Он равен интегралу по электронным координатам r . Приближение, которое позволяет существенно упростить рассмотрение, предполагая слабую зависимость электронного матричного элемента \mathbf{d}_{fi} от \mathcal{R} , называется приближением Кондона [98]. В этом случае $\mathbf{d}_{fi}(\mathcal{R}_0)$ можно вынести за интеграл по \mathcal{R} , и мнимая часть диэлектрической проницаемости ε_2 может быть выражена через так называемую функцию формы линии $G(\omega)$:

$$\varepsilon_2(\omega) = 4\pi^2 N \hbar^{-1} \left| \mathbf{e} \mathbf{d}_{fi} \right|^2 G(\omega), \quad (13.2)$$

где

$$G(\omega) = \left\langle \left\langle \sum_{\{n_f\}} \left| \int d\mathcal{R} \chi_f^*(\mathcal{R}) \chi_i(\mathcal{R}) \right|^2 \delta(\omega - \omega_{f\{n_f\}, i\{n_i\}}) \right\rangle \right\rangle_T.$$

Рассмотрим интеграл от $G(\omega)$ по всем частотам:

$$\begin{aligned} \int_0^\infty d\omega G(\omega) &= \\ &= \left\langle \left\langle \sum_{\{n_f\}} \int \int d\mathcal{R} d\mathcal{R}' \chi_f^*(\mathcal{R}) \chi_i(\mathcal{R}) \chi_f(\mathcal{R}') \chi_i^*(\mathcal{R}') \right\rangle \right\rangle_T. \end{aligned} \quad (13.3)$$

Пользуясь полнотой системы фоновых функций и их нормировкой как в конечном, так и в начальном состоянии, выражение (13.3) можно привести к виду

$$\begin{aligned} \int_0^\infty d\omega G(\omega) &= \left\langle \left\langle \int \int d\mathcal{R} d\mathcal{R}' \chi_i(\mathcal{R}) \chi_i^*(\mathcal{R}') \delta(\mathcal{R} - \mathcal{R}') \right\rangle \right\rangle_T = \\ &= \left\langle \left\langle \int d\mathcal{R} |\chi_i(\mathcal{R})|^2 \right\rangle \right\rangle_T = \langle \langle 1 \rangle \rangle_T = 1. \end{aligned} \quad (13.4)$$

Таким образом, введенная в (13.2) функция формы линии нормирована на единицу. Заметим, что в адиабатическом приближении и приближении Кондона вся зависимость от внешних параметров, например от температуры, заключена в $G(\omega)$. Поэтому равенство (13.4) приводит к важному закону постоянства площади линии поглощения:

$$\int_0^\infty \varepsilon_2(\omega) d\omega = \text{const}.$$

Это равенство можно использовать для проверки точности приближения Кондона.

Помимо функции формы линии $G(\omega)$ в теоретических исследованиях большое значение имеет ее Фурье-образ — так называемая характеристическая функция

$$g(t) = \int_{-\infty}^{\infty} d\omega e^{-i\omega t} G(\omega).$$

Используя известное выражение для Фурье-образа дельта-функции, можно показать, что

$$\begin{aligned}
g(t) &= \left\langle \left\langle \sum_{\{n_f\}} \left| \langle \{n_f\} | \{n_i\} \rangle \right|^2 \exp \left[\frac{i}{\hbar} \left(E_{f\{n_f\}} - E_{i\{n_i\}} \right) t \right] \right\rangle \right\rangle_T = \\
&= \left\langle \left\langle \sum_{\{n_f\}} \langle \{n_i\} | \{n_f\} \rangle e^{\frac{i}{\hbar} E_{f\{n_f\}} t} \langle \{n_f\} | \{n_i\} \rangle e^{-\frac{i}{\hbar} E_{i\{n_i\}} t} \right\rangle \right\rangle_T = \\
&= \left\langle \left\langle \sum_{\{n_f\}} \langle \{n_i\} | \{n_f\} \rangle \exp \left[\frac{i}{\hbar} \hat{\mathcal{H}}_{ph}^f t \right] \langle \{n_f\} | \{n_i\} \rangle \exp \left[-\frac{i}{\hbar} \hat{\mathcal{H}}_{ph}^i t \right] \right\rangle \right\rangle_T = \\
&= \left\langle \left\langle \sum_{\{n_f\}} \langle \{n_i\} | \exp \left[\frac{i}{\hbar} \hat{\mathcal{H}}_{ph}^f t \right] | \{n_f\} \rangle \langle \{n_f\} | \exp \left[-\frac{i}{\hbar} \hat{\mathcal{H}}_{ph}^i t \right] | \{n_i\} \rangle \right\rangle \right\rangle_T = \\
&= \left\langle \left\langle \langle \{n_i\} | \exp \left[\frac{i}{\hbar} \hat{\mathcal{H}}_{ph}^f t - \frac{i}{\hbar} \hat{\mathcal{H}}_{ph}^i t \right] | \{n_f\} \rangle \right\rangle \right\rangle_T. \quad (13.5)
\end{aligned}$$

Здесь введены операторы Гамильтона (12.6) для фоновых состояний в начальном и конечном электронных состояниях. При преобразованиях использовалась также полнота системы фоновых функций. Выражение (13.5) для характеристической функции по происхождению аналогично выражению, полученному при рассмотрении эффекта Оже и, в частности, не содержит волновой функции конечного состояния.

В случае линейной электрон-решеточной связи волновые функции каждой моды в основном и возбужденном состояниях отличаются лишь сдвигом положения равновесия. Воспользуемся равенством (12.8) и первым преобразованием, примененным в цепочке выражений (13.5):

$$\begin{aligned}
g(t) &= \left\langle \left\langle \sum_{\{n_f\}} \prod_{\mathbf{k}_s} \left| \int dQ_{\mathbf{k}_s} \chi_{n_{\mathbf{k}_s}^*}^*(Q_{\mathbf{k}_s} - \bar{Q}_{\mathbf{k}_s}^f) \chi_{n_{\mathbf{k}_s}^i}(Q_{\mathbf{k}_s} - \bar{Q}_{\mathbf{k}_s}^i) \right|^2 \times \right. \right. \\
&\quad \left. \left. \times \exp \left[\frac{i}{\hbar} E_{fi}^0 t + i(n_{\mathbf{k}_s}^f - n_{\mathbf{k}_s}^i) \right] \right\rangle \right\rangle_T. \quad (13.6)
\end{aligned}$$

Здесь $\chi_n(Q)$ — n -я собственная функция одномерного осциллятора с частотой $\Omega_s(\mathbf{k})$ в нормальных координатах, $n_{\mathbf{k}_s}^f$ и

$n_{\mathbf{k}s}^i$ — числа заполнения для моды $\mathbf{k}s$ из наборов $\{n_f\}$ и $\{n_i\}$, соответственно. Энергия чисто электронного перехода E_{fi}^0 равна $E_f^0 - E_i^0$. Принимая во внимание явный вид функций осциллятора $\chi_n(Q)$ (см., например, [49]), легко показать, что интеграл по $Q_{\mathbf{k}s}$ в (13.6) зависит от только от номеров $n_{\mathbf{k}s}^f$ и $n_{\mathbf{k}s}^i$ и от безразмерного фактора Хуанга — Риса $S_{\mathbf{k}s}$ (12.9).

Если предполагать, что локальных фононных мод нет, то взаимодействие с каждой отдельной модой мало (порядка N^{-1} , где N — число атомов в основной области кристалла), и факторы Хуанга — Риса также малы ($S_{\mathbf{k}s} \sim N^{-1}$). В этом случае, воспользовавшись рекуррентными соотношениями для функций осциллятора, можно показать, что основной вклад в интегралы перекрытия фононных функций вносят интегралы с малым изменением квантовых чисел осцилляторов. Используя рекуррентные соотношения для функций осциллятора, с точностью до линейных по S членов можно получить

$$\begin{aligned} \left| \int dQ \chi_n^*(Q - \bar{Q}^f) \chi_n(Q - \bar{Q}^i) \right|^2 &= 1 - (2n + 1)S, \\ \left| \int dQ \chi_{n\pm 1}^*(Q - \bar{Q}^f) \chi_n(Q - \bar{Q}^i) \right|^2 &= (2n + 1 \pm 1)S/2. \end{aligned}$$

Остальные интегралы (с изменением n более чем на единицу) имеют более высокий порядок малости по S . В таком случае сумма по $\{n_f\}$ сводится к трем членам для каждой моды:

$$\begin{aligned} g(t) = e^{iE_{fi}^0 t/\hbar} \left\langle \left\langle \prod_{\mathbf{k}s} [1 - (2n_{\mathbf{k}s}^i + 1)S_{\mathbf{k}s}(1 - \cos\Omega_s(\mathbf{k})t) + \right. \right. & \quad (13.7) \\ \left. \left. + iS_{\mathbf{k}s} \sin\Omega_s(\mathbf{k})t \right] \right\rangle_T. \end{aligned}$$

Поскольку каждое $n_{\mathbf{k}s}^i$ входит линейным образом, усреднение по начальному состоянию выполняется тривиально и сводится к замене $n_{\mathbf{k}s}^i$ на равновесные значения чисел заполнения соответствующей моды $\overline{n_{\mathbf{k}s}^i} = \left[e^{\hbar\Omega_s(\mathbf{k})/k_B T} - 1 \right]^{-1}$. Обратим внимание, что $2\overline{n_{\mathbf{k}s}^i} + 1 = \text{cth}(\hbar\Omega_s(\mathbf{k})/2k_B T)$. Поскольку отличие каждого из сомножителей от единицы мало, можно

воспользоваться тем, что $e^x \approx 1 + x$ для малых x , и перейти от произведения к экспоненте от суммы:

$$g(t) = e^{iE_{fi}^0 t/\hbar} \times \exp \left\{ - \sum_{\mathbf{k}s} S_{\mathbf{k}s} \left[(1 - \cos \Omega_s(\mathbf{k})t) \operatorname{cth} \frac{\hbar \Omega_s(\mathbf{k})}{2k_B T} - i \sin \Omega_s(\mathbf{k})t \right] \right\}. \quad (13.8)$$

Дальнейшее рассмотрение требует предположений о величине не самих факторов Хуанга — Риса, а их сумм по всем модам.

13.2 Слабая электрон-фононная связь

В ряде случаев взаимодействие электронов с фононами является настолько слабым, что $\sum_{\mathbf{k}s} S_{\mathbf{k}s} \ll 1$. При этом можно разложить экспоненту в (13.8), после чего получим

$$g(t) = e^{iE_{fi}^0 t/\hbar} \left\{ 1 + \sum_{\mathbf{k}s} S_{\mathbf{k}s} \left[- \operatorname{cth} \frac{\hbar \Omega_s(\mathbf{k})}{2k_B T} + \frac{\exp(-i\Omega_s(\mathbf{k})t)}{\exp(\hbar \Omega_s(\mathbf{k})/k_B T) - 1} + \frac{\exp(i\Omega_s(\mathbf{k})t)}{1 - \exp(-\hbar \Omega_s(\mathbf{k})/k_B T)} \right] \right\}.$$

Переходя к функции формы линии $G(\omega)$, получим

$$G(\omega) = \left[1 - \sum_{\mathbf{k}s} S_{\mathbf{k}s} \operatorname{cth} \frac{\hbar \Omega_s(\mathbf{k})}{2k_B T} \right] \delta(\omega - E_{fi}^0/\hbar) + \sum_{\mathbf{k}s} S_{\mathbf{k}s} \left[\frac{\delta(\omega - E_{fi}^0/\hbar + \Omega_s(\mathbf{k}))}{\exp(\hbar \Omega_s(\mathbf{k})/k_B T) - 1} + \frac{\delta(\omega - E_{fi}^0/\hbar - \Omega_s(\mathbf{k}))}{1 - \exp(-\hbar \Omega_s(\mathbf{k})/k_B T)} \right].$$

Этот результат допускает простую интерпретацию. Первый член соответствует чисто электронному переходу с энергией E_{fi}^0 без изменения фононных состояний. При увеличении температуры интенсивность этой бесфононной линии уменьшается. Второй и третий члены описывают поглощение с энергией, меньшей или большей энергии бесфононного перехода, соответственно. При этом одновременно происходит поглощение или испускание фонона с частотой $|\omega - E_{fi}^0/\hbar| = \Omega_s(\mathbf{k})$. Поскольку фононы непрерывно

распределены в некоторой области частот с плотностью состояний $\rho(\Omega)$, суммирование по этим состояниям можно заменить интегрированием, и мы получим

$$G(\omega) = \left[1 - \int_0^\infty d\Omega \overline{S\rho}(\Omega) \operatorname{cth} \frac{\hbar\Omega}{2k_B T} \right] \delta(\omega - E_{fi}^0/\hbar) + \\ + \frac{\overline{S\rho}(E_{fi}^0/\hbar - \omega) \theta(E_{fi}^0/\hbar - \omega)}{\exp[(E_{fi}^0 - \hbar\omega)/k_B T] - 1} + \frac{\overline{S\rho}(\omega - E_{fi}^0/\hbar) \theta(\omega - E_{fi}^0/\hbar)}{1 - \exp[(E_{fi}^0 - \hbar\omega)/k_B T]}.$$

В этом выражении третий член называется стоксовой линией поглощения, и он отличен от нуля для частот, больших частоты бесфононного перехода. Как видно, он пропорционален среднему по всем фононам с частотой Ω от произведения плотности фононных состояний на фактор Хуанга—Риса $\overline{S\rho}(\Omega)$. Поэтому все особенности плотности состояний фононов $\rho(\Omega)$ проявляются и в спектре оптического поглощения. Второй член описывает так называемую антистоксову линию поглощения, которая лежит по энергии ниже энергии бесфононного перехода. Антистоксова линия много слабее стоксовой, особенно для низких температур:

$$G(E_{fi}^0/\hbar + \Delta\omega) = e^{\hbar\Delta\omega/k_B T} G(E_{fi}^0/\hbar - \Delta\omega),$$

однако с ростом температуры различие в интенсивностях уменьшается. На рис. 33, *a* приведен спектр поглощения в случае слабой электрон-решеточной связи.

При усилении фактора Хуанга—Риса, когда сумма по всем модам становится порядка единицы, в разложении экспоненты в ряд становится невозможным ограничиться только первыми членами. При этом начинают появляться члены второго и более высоких порядков, связанные с поглощением или испусканием нескольких фононов: $\omega = E_{fi}^0 \pm \Omega_{s_1}(\mathbf{k}_1) \pm \Omega_{s_2}(\mathbf{k}_2) \pm \dots$

Возникающие особенности в

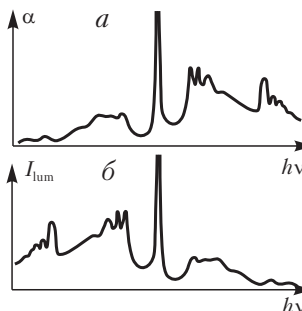


Рис. 33. Форма линии поглощения (*a*) и испускания (*б*) люминесценции в случае слабой электрон-фононной связи.

спектре поглощения могут быть связаны с наложением плотностей состояний комбинирующихся фононов.

Вообще говоря, характеристическая функция (13.7) является произведением по всем модам. Поэтому ее Фурье-образ является многократной сверткой всех возможных форм линий, каждая из которых связана с одной фононной модой. Этот результат является общим для случая невырожденных электронных переходов. Например, общую форму линии при включении дополнительного взаимодействия можно получить сверткой формы линии при учете только этого взаимодействия с прежней формой линии. При выполнении этой свертки необходимо центрировать все линии относительно бесфононной линии.

Можно сделать еще одно замечание общего характера о спектре люминесценции дефекта. Эта люминесценция связана с переходом из возбужденного состояния в основное. Выше для описания начального состояния предполагалось лишь, что система фононов находится в тепловом равновесии. Время жизни возбужденного электронного состояния относительно испускания фотона обычно много больше характерного времени фононной релаксации, приводящей к установлению теплового равновесия в фононной подсистеме. Поэтому при описании люминесценции можно предполагать, что происходит испускание фотона при переходе из термически равновесного по фононам возбужденного состояния в основное. Выражение для формы линии испускания будет совпадать с выражением для формы линии поглощения с той разницей, что E_{fi}^0 будет иметь другой знак. Отсюда можно сделать вывод, что в случае линейной электрон-решеточной связи форма линии поглощения является зеркальным отражением относительно бесфононной линии формы линии испускания люминесценции возбужденного дефекта (рис. 33, б).

13.3 Сильная электрон-фононная связь

Характерные черты спектра поглощения в случае сильной электрон-фононной связи можно разобрать на простом примере взаимодействия с оптическими колебаниями [95]. При этом предполагается, что частота оптических фононов не зависит от волнового вектора: $\Omega_{\text{opt}}(\mathbf{k}) = \Omega_0$. При этом характеристическая функция взаимодействия с оптической модой (13.8) принимает вид

$$g(t) = \exp \left[\frac{iE_{fi}^0 t}{\hbar} - S(1 - \cos \Omega_0 t) \operatorname{cth} \frac{\hbar \Omega_0}{2k_B T} + iS \sin \Omega_0 t \right].$$

Здесь

$$S = \sum_{\mathbf{k}} S_{\text{opt}, \mathbf{k}}$$

— величина, которая может быть и не малой (однако каждая из $S_{\text{opt}, \mathbf{k}}$ мала, что позволяет использовать выражение (13.8)). Если температура мала, $k_B T \ll \hbar \Omega_0$, то выражение для $g(t)$ приобретает вид

$$g(t) = \exp \left[\frac{iE_{fi}^0 t}{\hbar} - S + S e^{i\Omega_0 t} \right] = \exp \left[\frac{iE_{fi}^0 t}{\hbar} - S \right] \sum_{n=0}^{\infty} \frac{S^n e^{in\Omega_0 t}}{n!}.$$

Форма линии может быть легко получена Фурье-преобразованием:

$$G(\omega) = \sum_{n=0}^{\infty} \frac{e^{-S} S^n}{n!} \delta(\omega - E_{fi}^0 / \hbar - n\Omega_0).$$

Как видно, спектр поглощения имеет вид большого числа распределенных по закону Пуассона дельта-функций. В действительности предположение о постоянстве частот $\Omega_{\text{opt}}(\mathbf{k})$ не является строгим, поэтому дельта-функции размываются, и спектр становится непрерывным. Используя для $n!$ в случае больших n формулу Стирлинга, можно получить оценку формы линии для $k_B T \ll \hbar \Omega_0$:

$$G(\omega) = \frac{1}{\sqrt{2\pi S \Omega_0}} \exp \left[-\frac{(\omega - E_{fi}^0 / \hbar - S \Omega_0)^2}{2S \Omega_0^2} \right], \quad (13.9)$$

т. е. гауссову линию поглощения шириной $\sqrt{2S \Omega_0}$, центр которой находится при энергии $E_{fi}^0 + S \hbar \Omega_0$. Форма линии при различных S при нулевой температуре показана на рис. 34.

Полученный результат легко может быть интерпретирован на языке конфигурационных координат. Все оптические моды вносят линейный вклад в гамильтониан (12.6). Обозначим через Q_1 некоторую линейную комбинацию нормальных

координат $B^{-1} \sum_{\mathbf{k}} B_{\mathbf{k}} Q_{\mathbf{k}}$

(предполагается, что в основном состоянии $\overline{Q_{\mathbf{k}}^i} = 0$). Тогда в гамильтониан будет входить только BQ_1 . Постоянная B выбирается из тех соображений, чтобы при введении координаты Q_1 в гамильтониан коэффициент при Q_1^2 был бы равен $\Omega_0^2/2$ (т.е. $B^2 = 2S\Omega_0^2$). Введенная

таким образом мода называется модой взаимодействия и не является нормальной модой системы. При этом потенциальная энергия может быть изображена для двух состояний электронной системы в виде двух парабол (рис. 35). В основном состоянии при нулевой температуре координата Q_1 распределена по Гауссу (вероятность нахождения системы в точке Q_1 из-за нулевых колебаний равна $\exp(-Q_1^2 \Omega_0^2 / 2\hbar\Omega_0)$). Из-за того, что оптический переход протекает за времена, существенно меньшие времен релаксации решетки, изменения координат, в частности Q_1 , не происходит (принцип Франка—Кондона [98, 99]). Тогда линия поглощения пропорциональна

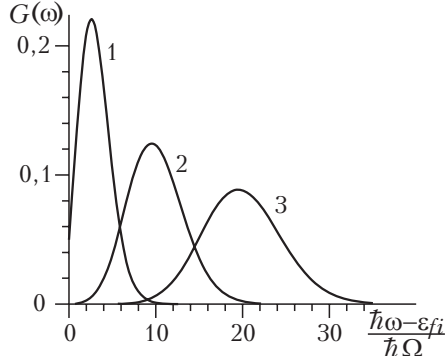


Рис. 34. Форма линии поглощения в случае сильной электрон-фононной связи: 1 — $S=3$; 2 — $S=10$; 3 — $S=20$. Энергия отсчитывается от энергии бесфононного перехода.

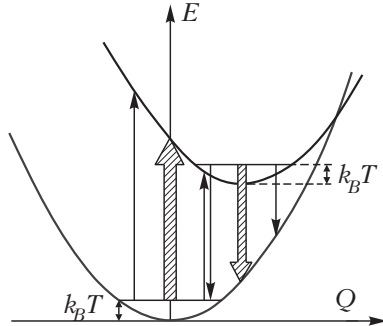


Рис. 35. Схема переходов между двумя уровнями при сильной электрон-фононной связи.

$$G(\omega) \sim \int dQ_1 \exp(-Q_1^2 \Omega_0 / 2\hbar) \delta(\hbar\omega - E_{fi}^0 + Q_1 \sqrt{2S\Omega_0} - S\hbar\Omega_0) \sim \\ \sim \exp \left[-\frac{(\omega - E_{fi}^0 / \hbar - S\Omega_0)^2}{2S\Omega_0^2} \right].$$

Тем самым получен результат, выведенный ранее более строго. Аналогичным образом можно рассмотреть и случай высоких температур. При этом начальное распределение по координате Q_1 определяется уже не нулевыми колебаниями, а температурным распределением Больцмана: $\exp(-Q_1^2 \Omega_0^2 / 2k_B T)$. Вследствие этого распределение поглощения имеет тоже гауссов вид, но с заменой $\hbar\Omega_0$ в дисперсии на $2k_B T$:

$$G(\omega) \sim \exp \left[-\frac{\hbar(\omega - E_{fi}^0 / \hbar - S\Omega_0)^2}{4S\Omega_0 k_B T} \right], \quad k_B T \gg \hbar\Omega_0. \quad (13.10)$$

В случае промежуточных температур, когда термические флуктуации сравнимы с квантовыми нулевыми колебаниями, $\hbar\Omega_0$ заменяется на $\hbar\Omega_0 \operatorname{cth}(\hbar\Omega_0 / 2k_B T)$:

$$G(\omega) \sim \exp \left[-\frac{(\omega - E_{fi}^0 / \hbar - S\Omega_0)^2}{2S\Omega_0^2 \operatorname{cth}(\hbar\Omega_0 / 2k_B T)} \right]. \quad (13.11)$$

Аналогичное рассмотрение можно провести и для испускания люминесценции. При этом в обоих предельных случаях получаются аналогичные (13.9) – (13.11) выражения, однако максимум сдвинут в точку $E_{fi}^0 / \hbar - S\Omega_0$. Тем самым линии поглощения и свечения оказываются зеркально-симметричными относительно энергии бесфононного перехода E_{fi}^0 / \hbar , а сдвиг между энергиями линий поглощения и излучения в случае сильной электрон-решеточной связи составляет $2S\Omega_0$ – удвоенный стоксов сдвиг для каждой из линий.

Данное рассмотрение взаимодействия электронных состояний с фононами для случая сильной связи справедливо не только для описания взаимодействия с оптическими фононами, но также и для случая взаимодействия с локальными колебаниями около дефекта. В этом случае Q_1 приобретает более очевидный физический смысл конфигурационной

координаты, описывающей смещение либо самого дефектного атома, либо ближайших к дефекту ионов.

Диаграмма конфигурационных координат оказывается чрезвычайно полезной при анализе сложных случаев с переходами между состояниями, хотя бы одно из которых вырождено. При этом, поскольку смещение из положения равновесия дефектного иона и/или окружающих ионов снимает вырождение, строгое решение задачи о форме линии становится громоздким, и удобно провести качественный анализ. Например, рассмотрим случай переходов из невырожденного основного состояния A в двукратно вырожденное возбужденное состояние E , взаимодействующее с двукратно вырожденной локальной фононной модой. Взаимодействие E -состояния с E -модой приводит к расщеплению этого уровня на два уровня, энергия которых зависит только от комбинации $\rho = \sqrt{Q_1^2 + Q_2^2}$ фононных конфигурационных координат Q_1 и Q_2 E -моды (рис. 36). В этих обозначениях распределение по ρ в начальном состоянии пропорционально

$$\rho \exp \left[-\frac{\rho^2 \Omega_0^2}{2\hbar\Omega_0 \operatorname{cth}(\hbar\Omega_0/2k_B T)} \right],$$

и его максимум сдвинут из точки $\rho = 0$. Легко видеть, что в этом случае линия поглощения будет иметь двугорбый характер из-за переходов на два уровня возбужденного расщепленного E -состояния. Можно отметить, что в случае, показанном на рис. 36, линия с меньшей энергией имеет большую ширину, чем линия с большей энергией (поскольку наклоны конфигурационных кривых начального состояния и

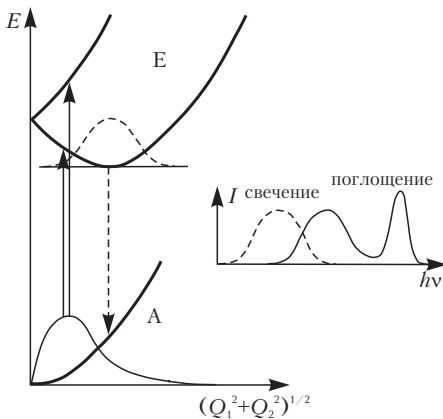


Рис. 36. Схема переходов в случае возникновения эффекта Яна-Теллера.

верхнего подуровня E -состояния ближе друг к другу).

Теорема о зеркальности линий поглощения и люминесценции больше не выполняется для вырожденных состояний. Действительно, в возбужденном электронном состоянии система перейдет в состояние метастабильного равновесия по фоновой подсистеме вблизи минимума нижней из двух кривых. Тем самым линия люминесценции будет одnogорбой и будет соответствовать переходам только с одного из возбужденных уровней.

13.4 Безызлучательные переходы

Кроме переходов из одного электронного состояния в другое путем поглощения или испускания фотона, в системе «дефект + фононы» возможны переходы без участия фотонов. В частности, таковыми являются безызлучательные переходы из возбужденного электронного состояния в основное.

Если бы адиабатическое приближение являлось точным, такие переходы были бы запрещены. Напомним, что для обоснования адиабатического приближения в (12.4) были опущены члены в фигурных скобках {...}. Если предполагается, что электронная функция $\psi(r; \mathcal{R})$ все-таки зависит от координат ионной подсистемы, эти члены приводят к связыванию различных состояний электронной подсистемы.

При рассмотрении уровней энергии в системе «дефект + решетка» (рис. 26) можно ввести две величины, полезные для последующего изложения: энергию бесфононного перехода E_{fi}^0 и энергию активации E_{fi}^A , равную расстоянию по шкале энергии между точкой пересечения конфигурационных кривых и минимумом возбужденного состояния. Если предполагать применимость гармонического приближения для всех Q , то

$$E_{fi}^A = (E_{fi}^0 - \Delta E)^2 / 4\Delta E, \quad (13.12)$$

где $\Delta E = \hbar\Omega S$.

В приближении слабой связи из (13.12) следует, что E_{fi}^A очень большое и попадает в область, где применимость гармонического приближения заведомо нарушается. В этом случае вероятность безызлучательных переходов из возбужденного состояния в основное состояние оказывается порядка [100]

$$w \sim \exp \left\{ -\frac{E_{fi}^0}{\hbar\Omega_M} \left[\ln \frac{E_{fi}^0}{\hbar\Omega_M (\bar{n}_M + 1)} - 1 \right] \right\}, \quad (13.13)$$

где Ω_M — максимальная частота фонона, для которого фактор Хуанга-Риса S_M отличен от нуля, \bar{n}_M — соответствующее число заполнения этой фононной моды. Из (13.13) видно, что вероятность безызлучательных переходов в этом приближении быстро падает с ростом энергии безызлучательного перехода E_{fi}^0 . Аналогичная зависимость была получена в работе [101]. В работе [102] было показано для редкоземельных ионов в ионном кристалле эта вероятность существенна (по сравнению с вероятностью излучательных переходов) только для $E_{fi}^0/\hbar\Omega_M \leq 6$. Типичные значения выражения в квадратных скобках в (13.13) больше или порядка двойки.

В противоположном пределе сильной связи электронов с фононами стоксов сдвиг ΔE уже не является малым, и в соответствии с формулой (13.12) точка пересечения двух конфигурационных кривых приближается к минимуму верхней конфигурационной кривой. В этом пределе вероятность безызлучательного перехода может быть записана [96, 101] как

$$w \sim \exp \left(-E_{fi}^A / \left(\frac{\hbar\bar{\omega}}{2} \operatorname{cth} \frac{\hbar\bar{\omega}}{2k_B T} \right) \right).$$

В пределе высоких температур это дает

$$w \sim \exp \left(-E_{fi}^A / k_B T \right),$$

а для низких —

$$w \sim \exp \left(-E_{fi}^A / \hbar\bar{\omega} \right).$$

Здесь $\bar{\omega}$ означает некоторую среднюю частоту фононов, на которой происходит взаимодействие. Этот результат типичен для активационных процессов с энергией активации E_{fi}^A , и связан с тем, что матричный элемент оператора неадиабатичности максимален в точке пересечения конфигурационных кривых. При низких температурах возможно туннелирование системы с одной конфигурационной кривой на другую.

§14 Поляроны

14.1 Поляроны большого радиуса. Деформационное электрон-фононное взаимодействие

Для исследования взаимодействия зонных электронов с фононам необходимо оценить коэффициенты $V_{\mathbf{k}s}$ в гамильтониане (12.6). Как будет видно из дальнейшего, наибольшую роль при взаимодействии с зонными состояниями играют длинноволновые фононы. Такого рода состояниями удобно изучать, пользуясь так называемым континуальным приближением, когда кристалл рассматривается как непрерывная среда. При этом акустические и оптические фононы описываются по-разному.

Как выше указывалось, акустические фононы приводят к изменению объема элементарной ячейки. Относительное изменение объема составляет $1 + \text{div } \mathbf{u}(\mathbf{r})$, где $\mathbf{u}(\mathbf{r})$ — смещение точки \mathbf{r} кристалла при акустической деформации (см. (12.5)). Изменение объема эквивалентно изменению постоянной решетки кристалла. Как видно из рис. 10, изменение расстояния между атомами приводит к сдвигу энергетических уровней. В простейшем приближении закон дисперсии для электронов в зоне s сдвигается на некоторую величину:

$$E_s(\mathbf{k}, \mathbf{u}(\mathbf{r})) = E_s(\mathbf{k}) - \sigma_s \text{div } \mathbf{u}(\mathbf{r}). \quad (14.1)$$

Последнее слагаемое в этой формуле называют деформационным потенциалом Бардина и Шокли [103]. Точные вычисления σ_s достаточно сложны. Оценка для σ_s дает значения порядка нескольких электрон-вольт (в металлах σ_s порядка энергии Ферми E_F). Коэффициент σ_s предполагается независимым от волнового вектора \mathbf{k} . Это справедливо для электронов низких энергий.

Выражение (14.1) справедливо лишь для изотропных кубических кристаллов. В более сложных системах помимо \mathbf{u} необходимо учитывать другие комбинации производных со

своими значениями постоянных коэффициентов. Формула (12.4) позволяет выразить значения деформационного потенциала через нормальные координаты фононов или через операторы рождения и уничтожения акустических фононов с помощью уравнений (12.5) и (12.7) (см. [1]):

$$-\sigma_s \operatorname{div} \mathbf{u}(\mathbf{r}) = N^{-1/2} \sum_{\mathbf{q}} F_{\text{ac}}^s(\mathbf{q}) (b_{\mathbf{q}LA} - b_{-\mathbf{q}LA}^+) e^{i\mathbf{q}\mathbf{r}} .$$

Здесь суммирование проводится по волновым векторам \mathbf{q} продольных акустических фононов. В кубических кристаллах поперечные фононы не должны учитываться в деформационном потенциале, поскольку поперечные фононы в изотропном кристалле не приводят к изменениям объема. Эти типы акустических фононов должны учитываться только в пьезоэлектрических и сегнетоэлектрических кристаллах (см. ниже, раздел 18.3). Коэффициент F_{ac}^s определяется выражением

$$F_{\text{ac}}^s(\mathbf{q}) = -i\sigma_s \sqrt{\frac{\hbar|\mathbf{q}|}{2Mc_{LA}}} , \quad (14.2)$$

где M — суммарная масса атомов в элементарной ячейке, c_{LA} — скорость звука в кристалле, связанная с частотой акустических фононов соотношением $\Omega_{LA}(\mathbf{q}) = c_{LA}|\mathbf{q}|$.

После перехода к электронным операторам рождения и уничтожения оператор электрон-фононного взаимодействия записывается в виде

$$\hat{\mathcal{H}}_{\text{int}} = N^{-1/2} \sum_{s\mathbf{q}} F_{\text{ac}}^s(\mathbf{q}) a_{\mathbf{k}+\mathbf{q},s}^+ a_{\mathbf{k}s} (b_{\mathbf{q}LA} - b_{-\mathbf{q}LA}^+) . \quad (14.3)$$

Каждое слагаемое в этой сумме соответствует рассеянию электрона с импульсом $\hbar\mathbf{k}$ в состояние с импульсом $\hbar(\mathbf{k} + \mathbf{q})$, сопровождающееся поглощением фонона с импульсом $\hbar\mathbf{q}$ или испусканием фонона с импульсом $-\hbar\mathbf{q}$. Переходами электрона между различными ветвями электронного спектра мы здесь пренебрегаем.

Для того чтобы оценить влияние фононов на электронный спектр, можно вычислить при нулевой температуре энергию электронного состояния во втором порядке теории возмущений:

$$E^{(2)}(\mathbf{k}) = E^{(0)}(\mathbf{k}) - \sum_{\mathbf{q}} \frac{\left| \langle \mathbf{k} - \mathbf{q}; 1_{\mathbf{q}} | \hat{\mathcal{H}}_{\text{int}} | \mathbf{k}; 0_{\mathbf{q}} \rangle \right|^2}{E^{(0)}(\mathbf{k} - \mathbf{q}) + \hbar\Omega_{LA}(\mathbf{q}) - E^{(0)}(\mathbf{k})}. \quad (14.4)$$

Матричный элемент гамильтониана взаимодействия между состоянием без фононов с волновым вектором $\mathbf{q} | \mathbf{k}; 0_{\mathbf{q}} \rangle$ и состоянием с одним фононом этой моды $|\mathbf{k} - \mathbf{q}; 1_{\mathbf{q}} \rangle$ может быть вычислен с помощью (14.3):

$$\left| \langle \mathbf{k} - \mathbf{q}; 1_{\mathbf{q}} | \hat{\mathcal{H}}_{\text{int}} | \mathbf{k}; 0_{\mathbf{q}} \rangle \right|^2 = \sigma_s^2 \frac{\hbar|\mathbf{q}|}{2MNc_{LA}}.$$

Знаменатель (14.4) можно разложить в ряд по \mathbf{k} . После подстановки $E^{(0)}$ в приближении эффективной массы m_s^* знаменатель (14.4) может быть представлен в виде

$$\hbar^2 [\mathbf{q}(\mathbf{q} + \mathbf{Q}) - 2\mathbf{k}\mathbf{q}] / 2m_s^*, \quad (14.5)$$

где $\mathbf{Q} = 2m_s^* c_{LA} \mathbf{q} / \hbar|\mathbf{q}|$. Переходя от суммирования по \mathbf{q} к интегрированию, несложно получить закон дисперсии электрона при малых \mathbf{k} :

$$E^{(2)}(\mathbf{k}) = \left(\frac{\hbar^2}{2m_s^*} - B \right) k^2 - A, \quad (14.6)$$

$$\text{где } A = \zeta \left[\frac{x^2}{2} - x + \ln(1+x) \right], \quad B = \frac{2\zeta}{3} \left[\frac{2x+1}{2(x+1)^2} - \frac{1}{2} + \ln(1+x) \right],$$

$x = \pi/aQ$, $\zeta = (\sigma_s m_s^* / \pi\hbar)^2 vQ/M$, v — объем элементарной ячейки кристалла, a — постоянная решетки. Учитывая, что A и B — положительные константы, можно заключить, что взаимодействие с фононами приводит к понижению дна зоны проводимости, увеличению эффективной массы электрона за счет вовлечения ионной подсистемы в движение электрона. В этих же приближениях можно оценить среднее число фононов, окружающих электрон:

$$\langle n_{\mathbf{q}} \rangle = \zeta [\ln(1+x) - x/(1+x)] / \hbar Q c_{LA}.$$

В случае слабой связи, когда только и можно пользоваться теорией возмущения, $\langle n_{\mathbf{q}} \rangle$ оказывается меньше единицы. Такое электронное состояние, окруженное «облаком» виртуальных

фононов, называется поляроном. Это название используется для состояний, являющихся результатом взаимодействия с любым типом фононов, независимо от наличия или отсутствия вызванной фононами поляризации решетки.

Радиус такого полярона можно оценить из соотношения неопределенности для волнового вектора виртуальных фононов, значение которого по модулю порядка Q . Тем самым решетка возмущается электроном на расстояниях порядка $2\pi/Q$, что в рассматриваемом случае много больше постоянной решетки a . В связи с этим такие поляроны называются поляронами большого радиуса.

Уравнение (14.6) справедливо только для $|\mathbf{k}| < Q/2$. Для больших \mathbf{k} знаменатель (14.5) обращается в нуль при некоторых значениях \mathbf{q} . При этом становится возможным процесс реального испускания фонона электроном. Условие $|\mathbf{k}| < Q/2$ эквивалентно условию черенковского излучения фотона электроном в среде. Скорость электрона $\hbar|\mathbf{k}|/m_s^*$ для порогового значения $|\mathbf{k}| = Q/2$ равна скорости звука в среде c_{LA} , поэтому только электроны с большей скоростью могут испускать звуковые волны – фононы. Вводя в знаменатель формулы (14.4) согласно адиабатическому приближению теории возмущений бесконечно малую мнимую часть, получим, что энергия $E^{(2)}(\mathbf{k})$ также становится комплексной:

$$E^{(2)}(\mathbf{k}) = E^{(0)}(\mathbf{k}) + \Delta(\mathbf{k}) + i\Gamma(\mathbf{k}),$$

где $\Delta(\mathbf{k})$ определяет сдвиг электронного уровня за счет взаимодействия с фононами, а $\Gamma(\mathbf{k})$ – его ширину:

$$\begin{aligned} \Delta(\mathbf{k}) &= -V.p. \sum_{\mathbf{q}} \frac{\left| \langle \mathbf{k} - \mathbf{q}; 1_{\mathbf{q}} | \hat{\mathcal{H}}_{\text{int}} | \mathbf{k}; 0_{\mathbf{q}} \rangle \right|^2}{E^{(0)}(\mathbf{k} - \mathbf{q}) + \hbar\Omega_{LA}(\mathbf{q}) - E^{(0)}(\mathbf{k})}, \\ \Gamma(\mathbf{k}) &= \pi \sum_{\mathbf{q}} \left| \langle \mathbf{k} - \mathbf{q}; 1_{\mathbf{q}} | \hat{\mathcal{H}}_{\text{int}} | \mathbf{k}; 0_{\mathbf{q}} \rangle \right|^2 \times \\ &\quad \times \delta(E^{(0)}(\mathbf{k} - \mathbf{q}) + \hbar\Omega_{LA}(\mathbf{q}) - E^{(0)}(\mathbf{k})). \end{aligned} \quad (14.7)$$

Таким образом, электронные состояния являются стационарными лишь для нулевой температуры и малых волновых векторов $|\mathbf{k}| < Q/2$, а для больших волновых векторов, и тем самым, больших энергий, становятся нестационарными.

Типичные времена жизни таких состояний порядка $\hbar/2\Gamma \approx 10^{-13} \div 10^{-12}$ с.

Для конечных температур необходимо учесть равновесное распределение фононов и включить не только процессы испускания фононов, но и их поглощения:

$$\Delta(\mathbf{k}) = -\sum_{\pm} \left\langle \left\langle \text{V.p.} \sum_{\mathbf{q}} \frac{\left| \langle \mathbf{k} \mp \mathbf{q}; n_{\mathbf{q}} \pm 1 | \hat{\mathcal{H}}_{\text{int}} | \mathbf{k}; n_{\mathbf{q}} \rangle \right|^2}{E^{(0)}(\mathbf{k} \mp \mathbf{q}) \pm \hbar\Omega_{LA}(\mathbf{q}) - E^{(0)}(\mathbf{k})} \right\rangle \right\rangle_T,$$

$$\Gamma(\mathbf{k}) = \pi \sum_{\pm} \left\langle \left\langle \sum_{\mathbf{q}} \left| \langle \mathbf{k} \mp \mathbf{q}; n_{\mathbf{q}} \pm 1 | \hat{\mathcal{H}}_{\text{int}} | \mathbf{k}; n_{\mathbf{q}} \rangle \right|^2 \times \right. \right. \\ \left. \left. \times \delta(E^{(0)}(\mathbf{k} \mp \mathbf{q}) \pm \hbar\Omega_{LA}(\mathbf{q}) - E^{(0)}(\mathbf{k})) \right\rangle \right\rangle_T.$$

Мы здесь опять использовали обозначение $\langle \langle \dots \rangle \rangle_T$ для усреднения по числам заполнения фононов $\{n_{\mathbf{q}}\}$. После выполнения этого усреднения ширина полярного состояния оказывается равной

$$\Gamma(\mathbf{k}) = \sum_{\mathbf{q}} \frac{\pi\sigma^2 \hbar |\mathbf{q}|}{2MNc_{LA}} \left[\frac{\delta(E^{(0)}(\mathbf{k} - \mathbf{q}) + \hbar\Omega_{LA}(\mathbf{q}) - E^{(0)}(\mathbf{k}))}{1 - \exp(-\hbar\Omega_{LA}(\mathbf{q})/k_B T)} + \right. \\ \left. + \frac{\delta(E^{(0)}(\mathbf{k} + \mathbf{q}) - \hbar\Omega_{LA}(\mathbf{q}) - E^{(0)}(\mathbf{k}))}{\exp(\hbar\Omega_{LA}(\mathbf{q})/k_B T) - 1} \right]. \quad (14.8)$$

При $T > 0$ это выражение отлично от нуля при всех \mathbf{k} , т.е. электронные состояния перестают быть стационарными за счет поглощения и испускания фононов. Второй член в (14.8) уменьшается с падением температуры, и для нулевых температур (14.8) переходит в (14.7).

14.2 Поляроны большого радиуса. Поляризационное взаимодействие

В ионных кристаллах большую роль играют оптические фононы, которые не приводят к изменению объема, но связаны с возникновением поляризации кристалла. В полную поляризацию кристалла вносят вклад как ионы, так и электроны. С оптическими колебаниями связана только ионная

составляющая поляризации. Для того, чтобы выделить вклад ионной подсистемы, заметим, что на низких частотах (в статическом случае $\omega \ll \Omega_{LO}$) в поляризацию вносят вклад как ионная, так и электронная подсистемы; если же частоты выбрать в области прозрачности кристалла ($\Omega_{LO} \ll \omega \ll E_g/\hbar$), то лишь вклад электронной подсистемы остается неизменным. Вклад ионной подсистемы в последнем диапазоне частот быстро уменьшается (как Ω_{LO}^2/ω^2), поскольку ионы не успевают следить за изменяющимся значением электрического поля. Таким образом, интересующая нас ионная часть поляризации равна

$$\mathbf{P}(\mathbf{r}) = \frac{1}{4\pi} \left[\left(1 - \frac{1}{\epsilon_0}\right) \mathbf{D}(\mathbf{r}) - \left(1 - \frac{1}{\epsilon_\infty}\right) \mathbf{D}(\mathbf{r}) \right] = \frac{\mathbf{D}(\mathbf{r})}{4\pi\epsilon^*},$$

где ϵ_0 и ϵ_∞ — соответственно статическая и высокочастотная диэлектрические проницаемости, а

$$\epsilon^* = \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right)^{-1} \quad (14.9)$$

— эффективная диэлектрическая проницаемость. Инерционная часть поляризации создает в кристалле электрическое поле, потенциал которого равен

$$4\pi \int \frac{\text{div } \mathbf{P}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r'. \quad (14.10)$$

В представлении вторичного квантования гамильтониан взаимодействия имеет вид (14.3) с заменой F_{ac}^s на

$$F_{\text{opt}}^s(\mathbf{q}) = -\frac{1}{|\mathbf{q}|} \sqrt{\frac{2\pi e^2 \hbar \Omega_{LO}}{v \epsilon^*}}. \quad (14.11)$$

Этот гамильтониан взаимодействия был впервые введен Фрелихом [104, 105].

Для характеристики силы электрон-фононной связи удобно ввести безразмерный параметр

$$\alpha = \frac{e^2}{\epsilon^*} \sqrt{\frac{m_s^*}{2\hbar^3 \Omega_{LO}}}.$$

Для большинства ионных кристаллов его значение изменяется в пределах от 1 до 10. Так же как и в случае взаимодействия с акустическими фононами, можно по формуле (14.4) оценить влияние фононов на сдвиг края зоны проводимости и изменения эффективной массы электрона. При $\alpha < 1$ справедлива теория возмущений, и можно получить

$$E^{(2)}(\mathbf{k}) = \frac{\hbar^2 k^2}{2m_s^*} \left(1 - \frac{\alpha}{6}\right) - \alpha \hbar \Omega_{LO} .$$

Среднее число виртуальных фононов, входящих в образование «одетого» электрона, оказывается равным $\alpha/2$. Аналогичные результаты были получены более сложным образом для промежуточной области $\alpha < 6$ без использования теории возмущений. Фейнман, используя метод интегрирования по траекториям, получил следующее выражение для энергии [106, 107]:

$$E(\mathbf{k}) = \frac{\hbar^2 k^2}{2m_s^* (1 + \alpha/6)} - \alpha \hbar \Omega_{LO} .$$

Необходимо отметить, что теория поляронов оказалась хорошей моделью для проверки и развития теории взаимодействия частицы с квантованными полями. Поэтому в развитии этой теории участвовали многие крупные физики.

Как уже обсуждалось в разделе 14.1, поляронные состояния с низкой энергией стабильны при нулевой температуре. Однако порог нестационарных состояний отличается от случая акустических фононов из-за независимости (в нулевом приближении) Ω_{LO} от \mathbf{q} . Именно, электрон может испустить реальный фонон, если его энергия превышает энергию продольного оптического фонона $\hbar \Omega_{LO}$. Ширина поляронного состояния равна

$$\Gamma(\mathbf{k}) = \sum_{\mathbf{q}} \frac{\pi}{N} \left| F_{\text{opt}}^s(\mathbf{q}) \right|^2 \left[\frac{\delta(E^{(0)}(\mathbf{k} - \mathbf{q}) + \hbar \Omega_{LO} - E^{(0)}(\mathbf{k}))}{1 - \exp(-\hbar \Omega_{LO}/k_B T)} + \frac{\delta(E^{(0)}(\mathbf{k} + \mathbf{q}) - \hbar \Omega_{LO} - E^{(0)}(\mathbf{k}))}{\exp(\hbar \Omega_{LO}/k_B T) - 1} \right] . \quad (14.12)$$

Процессы испускания и поглощения фононов подробно рассматриваются в следующем томе.

Пороговая энергия для испускания оптических фононов оказывается заметной ($\hbar\Omega_{LO} \sim 10 \div 40$ мэВ), в то время как пороговая энергия для испускания акустических фононов крайне мала (порядка $m_{ac}^* c_{ac}^2/2 \sim 10^{-5} \div 10^{-4}$ эВ).

14.3 Континуальное приближение

Описание электрон-фононного взаимодействия, проведенное в предыдущем разделе, основывалось на предположении об относительно слабом электрон-фононном взаимодействии. В настоящем разделе будет рассмотрен случай сильного взаимодействия. Такой случай не реализуется в большинстве кристаллов, но с качественной точки зрения очень важно.

Предположение, что взаимодействие между электроном и поляризацией решетки может привести к образованию потенциальной ямы и захвату электрона в эту яму, было впервые высказано Ландау [108] и реализовано в деталях Пекаром [109]. Пекар использовал для описания электрон-решеточного взаимодействия континуальное приближение, когда деформация и поляризация рассматриваются как непрерывные плавные функции, изменяющиеся на расстояниях, больших постоянной решетки. Отдельные фононы в таком подходе не рассматриваются. Такое приближение справедливо для сильной электрон-фононной связи ($\alpha \gg 1$).

Энергия электрона в кристалле может быть записана в виде функционала от волновой функции электрона $\psi(\mathbf{r})$ и ионной поляризации $\mathbf{P}(\mathbf{r})$:

$$E\{\psi, \mathbf{P}\} = \int \psi^*(\mathbf{r}) \left(-\frac{\hbar^2 \nabla^2}{2m^*} \right) \psi(\mathbf{r}) d^3r + \quad (14.13)$$

$$+ 2\pi\epsilon^* \int \mathbf{P}^2(\mathbf{r}) d^3r - \int \mathbf{P}(\mathbf{r}) \mathbf{D}(\mathbf{r}) d^3r,$$

где второе слагаемое — потенциальная энергия ионной поляризации кристалла:

$$U = \frac{1}{2} \int \mathbf{D}_1(\mathbf{r}) \mathbf{P}(\mathbf{r}) d^3r = 2\pi\epsilon^* \int \mathbf{P}^2(\mathbf{r}) d^3r,$$

\mathbf{D}_1 — вектор индукции, связанный с поляризацией $\mathbf{P}(\mathbf{r})$ (где $\mathbf{D}_1 = 4\pi\epsilon^* \mathbf{P}$),

$$\mathbf{D}(\mathbf{r}) = e \int |\psi(\mathbf{r}')|^2 \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} d^3 r' \quad (14.14)$$

— вектор индукции, создаваемый зарядом электрона с распределением $|\psi|^2$. Третье слагаемое в (14.13) определяет энергию взаимодействия электрона с решеткой. В таком подходе учитывается кинетическая энергия электрона (первое слагаемое в (14.13)), однако в соответствии с адиабатическим приближением кинетическая энергия решетки опускается.

Для нахождения ψ и \mathbf{P} необходимо минимизировать функционал (14.13) при условии нормировки волновой функции $\psi(\mathbf{r})$. В первую очередь минимизируем $E\{\psi, \mathbf{P}\}$ по отношению к вариациям $\mathbf{P}(\mathbf{r})$. Это может быть сделано путем решения уравнения на функциональную производную $\delta E\{\psi, \mathbf{P}\}/\delta \mathbf{P} = 0$. Решение этого уравнения дает связь $\mathbf{P}(\mathbf{r})$ с индукцией (14.14), создаваемой электроном:

$$\mathbf{P}(\mathbf{r}) = \mathbf{D}(\mathbf{r})/4\pi\epsilon^* \quad (14.15)$$

Подставляя это решение в $E\{\psi, \mathbf{P}\}$, находим функционал, зависящий уже только от волновой функции электрона ψ :

$$E\{\psi\} = \frac{\hbar^2}{2m^*} \int |\nabla\psi(\mathbf{r})|^2 d^3 r - \frac{1}{8\pi\epsilon^*} \int \mathbf{D}^2(\mathbf{r}) d^3 r,$$

где \mathbf{D} определено в (14.14).

Пекар [109] нашел минимум этого функционала с помощью пробной функции, зависящей от параметра r_0 :

$$\psi_0(\mathbf{r}) = \frac{1 + r/r_0}{\sqrt{7\pi r_0^3}} \exp(-r/r_0), \quad (14.16)$$

которая является сферически симметричной, нормированной и достаточно гладкой в точке $\mathbf{r} = 0$ (в этой точке имеет нулевую производную). Минимизация функции $E(r_0) = E\{\psi_0(\mathbf{r})\}$ относительно r_0 дает значение $r_0 = 2\hbar^2 \epsilon^*/m^* e^2$. Условием применимости континуального приближения является $r_0 \gg a$, где a — постоянная решетки. Значение энергии полярона для пробной функции (14.16) оказывается равным $E_0 = -0,054 m^* e^2/\hbar^2 \epsilon^*$. Эта энергия определяет энергию термической диссоциации полярона с разрушением

поляризационной потенциальной ямы. Энергия оптической диссоциации полярона E_1 больше E_0 в соответствии с принципом Франка-Кондона. Оптический переход происходит без разрушения потенциальной ямы. Это означает, что необходимо вычислить потенциальную яму (14.10), созданную поляризацией решетки (14.15), вызванной волновой функцией (14.16), и найти в такой потенциальной яме энергию основного состояния. Численный расчет дает величину $E_1 = -0,163 m^* e^2 / \hbar^2 \epsilon^*$, что в три раза больше E_0 . Поскольку при оптическом переходе яма сразу не разрушается, то абсолютное значение E_1 равно энергии фотоионизации полярона.

Взаимодействие электрона с оптическими колебаниями в ионных кристаллах приводит к возникновению полярона большого радиуса. Но в ряде кристаллов наблюдаются поляроны малого радиуса. Образование таких возбуждений также может быть рассмотрено в континуальном приближении. Акустические деформационные фононы играют в этом процессе важную роль [110, 111].

Деформационный потенциал может быть выражен через относительное изменение объема элементарной ячейки $\rho(\mathbf{r}) = \text{div } \mathbf{u}(\mathbf{r})$ в виде $-\sigma\rho(\mathbf{r})$ (14.1). Плотность энергии деформации равна $\beta\rho^2/2$, где β — упругая постоянная кристалла. Таким образом энергия поляронного состояния является функционалом $E\{\psi, \mathbf{P}, \rho\}$, который отличается от (14.13) наличием двух дополнительных членов

$$-\sigma \int |\psi|^2 \rho(\mathbf{r}) d^3 r + \frac{\beta}{2} \int \rho^2(\mathbf{r}) d^3 r .$$

Минимизация функционала $E\{\psi, \mathbf{P}, \rho\}$ по \mathbf{P} дает тот же результат (14.15), что и раньше, минимизация по ρ дает $\rho(\mathbf{r}) = \sigma|\psi|^2/\beta$. Подстановка этих значений в $E\{\psi, \mathbf{P}, \rho\}$ дает зависящий только от $\psi(\mathbf{r})$ функционал. Обсудим некоторые общие свойства этого функционала. Пробная функция может быть выбрана в виде

$$\psi(\mathbf{r}) = (\mu/a)^{3/2} f(\mu r/a) , \quad (14.17)$$

зависящем от вариационного параметра μ . Функция $f(x)$

предполагается нормированной: $4\pi \int_0^\infty x^2 f^2(x) dx = 1$. Если

среднее значение x для функции распределения $f^2(x)$ порядка единицы, то среднее значение r для распределения $|\psi|^2$ порядка a/μ . Поскольку для применимости континуального приближения a/μ должно быть больше a , то μ должно быть меньше единицы.

Подстановка (14.17) в функционал $E\{\psi\}$ и вынесение степеней μ за знаки интегралов дает

$$E(\mu) = B(\mu^2 - g_d \mu^3 - g_p \mu), \quad (14.18)$$

где B , g_d и g_p — константы, определяемые конкретным видом $f(x)$. Если выбрать $f(x) = 2^{3/4} \exp(-\pi x^2)$, то эти константы равны

$$B = \frac{3\pi\hbar^2}{2m^* a^2}, \quad g_d = \frac{\sigma m^*}{3\pi\beta\hbar^2}, \quad g_p = \frac{2e^2 m^* a}{3\pi\epsilon^* \hbar^2}.$$

Видно, что B связано с кинетической энергией электрона, g_d определяется деформационным взаимодействием с акустическими фононами, а g_p — поляризационным взаимодействием с оптическими фононами. Значения g_d и g_p определяют тип кривой $E(\mu)$. Экстремумы кривой $E(\mu)$ определяются уравнением

$$2\mu - 3g_d \mu^2 - g_p = 0, \quad (14.19)$$

решение которого имеет вид

$$\mu_{1,2} = [1 \pm \sqrt{1 - 3g_p g_d}] / 3g_d.$$

Полезно построить не только кривые $E(\mu)$ (рис. 37, а), но также и фазовую диаграмму в координатах силы электрон-фононной связи для деформационного и поляризационного взаимодействий (рис. 37, б). Значение μ_1 ($\mu_1 < \mu_2$) соответствует локальному минимуму кривой $E(\mu)$, а значение μ_2 — локальному максимуму. В случае $g_d \neq 0$ энергия $E(\mu)$ с ростом μ стремится к $-\infty$. Нас интересует только область применимости континуального приближения ($0 \leq \mu \leq 1$). В случае $\mu_1 < 1 < \mu_2$ (можно показать, что он реализуется при $g_p + 3g_d < 2$) у функционала $E(\mu)$ имеется только один минимум, соответствующий состоянию с относительно большим радиусом (так называемый свободный полярон, обозначаемый буквой F).

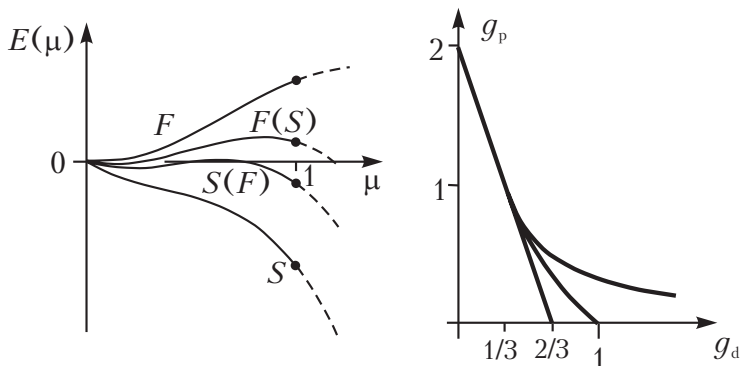


Рис. 37. Зависимость энергии $E(\mu)$ от вариационного параметра μ (левая панель) и фазовая диаграмма существования различных состояний полярона (правая панель).

Если уравнение (14.19) не имеет решений ($3g_d g_p > 1$) или $\mu_1 > 1$ ($3g_d < 1$, $g_p + 3g_d > 2$) энергия $E(\mu)$ уменьшается с ростом μ до тех пор, пока не станет нарушаться условие применимости континуального приближения ($\mu < 1$). Тем самым для этого случая нет вариационного решения задачи о минимизации функционала $E(\mu)$ в области $\mu < 1$, и для нахождения состояния автолокализованного полярона требуется микроскопический подход. Такое состояние обозначается буквой S . Континуальное приближение дает только тот результат, что нет решений с большим радиусом, и что стабильное состояние полярона имеет малый радиус. Энергия такого полярона может быть грубо оценена с использованием формулы (14.18) при $\mu = 1$.

В той части фазовой диаграммы, которая соответствует случаю $3g_d g_p < 1$, $3g_d > 1$, $g_p + 3g_d > 2$, одновременно могут сосуществовать оба типа решений. В зависимости от их относительных энергий могут реализовываться два случая: стабильное свободное решение и метастабильное автолокализованное (обозначаемое как $F(S)$), или наоборот ($S(F)$). Для случая $S(F)$ активационный барьер между двумя состояниями может быть оценен как

$$E^A = 4B(1 - 3g_p g_d)^{3/2} / 27g_d^2.$$

В противоположном случае $F(S)$ активационный барьер не

может быть оценен достаточно корректно, поскольку континуальное приближение уже плохо применимо для $\mu = 1$.

Подобный анализ был проведен в работах [112, 113] для более сложного случая автолокализации полярона в присутствии дефекта.

Автолокализация играет значительную роль при исследовании взаимодействия света с твердыми телами. При предыдущем рассмотрении обсуждалась автолокализация заряда без уточнения, электрон это или дырка. Чаще всего наблюдается автолокализация дырки, которая хорошо изучена в ионных кристаллах. Подробное описание структуры такой дырки представлено ниже.

В твердых телах могут наблюдаться более сложные случаи образования автолокализованных состояний. Автолокализованный экситон (self-trapped exciton, STE) является примером такого возбуждения. Число параметров при этом много больше, чем в случае отдельной частицы, поэтому поведение электронно-дырочной пары более сложное и разнообразное. В случае двух частиц с противоположным знаком заряда роль поляризационного взаимодействия эффективно ослабевает из-за дальнедействующего характера этого взаимодействия (поля электрона и дырки частично компенсируют друг друга). Поэтому можно рассматривать только короткодействующее деформационное взаимодействие частиц с решеткой. Основное влияние поляризации решетки имеет на изменение кулоновского взаимодействия (что выражается во введении постоянных ϵ_0 или ϵ_∞ , как это обсуждалось в разделе 6.4).

Здесь мы представим только пример фазовой диаграммы для отношения масс $m_e^*/m_h^* = 0,25$ и для случая пренебрежения поляризационным взаимодействием [114]. Фазовая диаграмма на рис. 38 показывает зависимость состояний от деформационных потенциалов g_d^e и g_d^h , отображаемых по осям абсцисс и ординат, соответственно. Предполагается, что $g_d^{e,h}$ могут принимать любой знак, и диаграмма имеет центр инверсии.

Для малых $|g_d^e|$ и $|g_d^h|$ стабильным является состояние свободного экситона F_{ex} , представляющего собой экситон Ванье-Мотта. С увеличением, скажем, силы взаимодействия g_d^h , происходит автолокализация дырки (S_h). Электрон в этом

случае связывается на этой дырке за счет кулоновского взаимодействия, аналогично случаю дефектного состояния большого радиуса. Такая конфигурация обозначается как $S_h:F_e$. В противоположном случае (большие g_d^e и относительно малые g_d^h) конфигурация меняется на обратную ($S_e:F_h$).

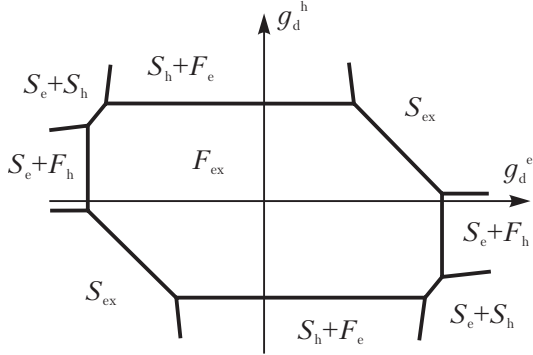


Рис. 38. Фазовая диаграмма для экситона на плоскости деформационных потенциалов g_d^e и g_d^h [114]

меняется на обратную ($S_e:F_h$).

Наиболее интересным случаем представляется случай относительно больших сил связи. Если g_d^e и g_d^h имеют одинаковый знак, возможно создание экситона малого радиуса (S_{ex}). Такой экситон может автолокализовываться даже в случаях, когда ни электрон, ни дырка не могут сделать этого по отдельности за счет конструктивного совмещения деформационного взаимодействия в том случае, если обе частицы сжимают (для положительных $g_d^{e,h}$) или растягивают (для отрицательных $g_d^{e,h}$) решетку. В том случае, если знаки этих постоянных противоположны, их совместное взаимодействие деструктивно, и в принципе может образовываться состояние из автолокализованного электрона и автолокализованной дырки, разделенных некоторым расстоянием ($S_e:S_h$).

Автолокализованные экситоны S_{ex} могут сосуществовать совместно со свободными F_{ex} . Эти два состояния разделены энергетическим барьером, как это описывалось для дырок (рис. 39). Наличие этого барьера обсуждалось Рашбой [115, 116], Тоязавой [110] и Моттом и Стоунхемом [117].

При предыдущем рассмотрении использовались параболические законы дисперсии для электронов и дырок с эффективными массами m_e^* и m_h^* . Для реальных кристаллов

необходимо принимать во внимание более сложные законы дисперсии. Кинетическая энергия для автолокализованного состояния может быть оценена равной половине ширины соответствующей зоны. Это дает простой энергетический критерий для существования автолокализованных состояний: такие состояния могут реализовываться, если сумма половины ширины зоны и энергии деформации и поляризации решетки меньше, чем энергия взаимодействия частицы с решеткой. Таким образом, автолокализованные состояния могут образовываться только в кристаллах с относительно узкими зонами. Это объясняет тот факт, что в щелочно-галогидных кристаллах дырки могут образовывать автолокализованные состояния, а электроны не могут.

Массы таких поляронов зависят от их радиуса. Масса полярона большого радиуса приблизительно равна массе свободной частицы, в то время как масса автолокализованного полярона много больше массы свободной частицы. Таким образом, полярон малого радиуса представляет собой очень тяжелую частицу, которая не может двигаться сквозь кристалл как когерентная частица. Для таких частиц понятие волнового вектора оказывается неприменимым.

14.4 Автолокализованные дырки и экситоны в щелочно-галогидных кристаллах. Образование дефектов

Как выше уже указывалось, с помощью континуального приближения невозможно описать микроскопическую структуру автолокализованного состояния. Эта структура сильно зависит от типа кристаллической решетки и может быть достаточно нетривиальной (прекрасный обзор автолокализованных экситонов представлен в [118]). В

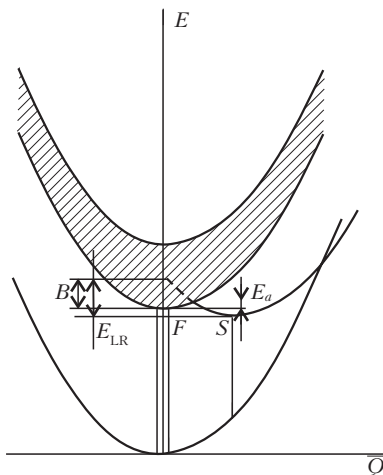


Рис. 39. Барьер между свободными и автолокализованными состояниями экситона.

качестве примера мы обсудим структуру автолокализованных дырок в типичном щелочно-галогидном кристалле со структурой поваренной соли – KCl.

Детальная структура таких автолокализованных дырок была получена с помощью экспериментов по электронному парамагнитному резонансу (ЭПР) (см., например, [119, 120]) и была подтверждена с помощью двойного электронно-ядерного резонанса (обзор применения этих методов можно найти, например, в [121]). Известно, что состояния дырок в валентной зоне порождены p -состояниями иона галоида ($3p$ состояниями Cl^-). Локализация дырки приводит к образованию в KCl двухузельной молекулы Cl_2^- , а не атомарного состояния Cl^0 . Такие центры получили название V_K -центров. Подобная галоидная молекула может быть ориентирована в направлении $\langle 110 \rangle$ в кристалле со структурой NaCl (ГЦК) и в направлении $\langle 100 \rangle$ в кристалле со структурой CsCl (ОЦК) (рис. 40). Расстояние между ионами галоидов в Cl_2^- оказывается существенно меньше, чем расстояния между галоидами в невозмущенном кристалле (уменьшение расстояния порядка 30–40%), поэтому оправданным представляется рассмотрение V_K -центра как молекулы в кристалле. Поэтому для классификации состояний V_K -центра можно использовать молекулярные обозначения. Основное состояние этой молекулы имеет конфигурацию $^2\Sigma_u^+$, в то время как возбужденные состояния – конфигурации $^2\Pi_g^-, ^2\Pi_u^-$ и $^2\Sigma_g^+$. Подробное описание свойств автолокализованных дырок можно найти в работах [122, 123].

Центры такого типа оказываются положительно заряженными, и они могут захватить электрон. После такого захвата нейтральное образование имеет электронную структуру молекулы Cl_2^- , окруженную электроном, размазанным по ближайшей конфигурационной сфере из щелочных ионов. В обозначениях предыдущего раздела такое состояние автолокализованного экситона можно описать как промежуточное между $S_h:F_e$ и S_{ex} -состояниями. После

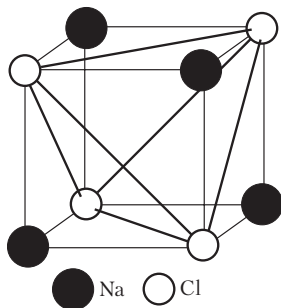


Рис. 40. Направления образования автолокализованного экситона в кристалле NaCl.

рекомбинации электрон переходит на молекулу Cl_2^- , в результате чего образуется основное состояние двух ионов Cl^- в решетке. Симметрия такого основного состояния — $^1\Sigma_g^+$, а симметрия молекулы Cl_2^- , вокруг которой находится электрон, может быть $^1\Sigma_u^+$ (синглетное состояние) или $^3\Sigma_u^+$ (триплетное состояние).

Поскольку электрон экранирует кулоновское поле дырки, потенциальная яма для автолокализованного экситона значительно более мелкая и плоская, чем для соответствующей автолокализованной дырки. Экситонная молекула может сдвигаться вдоль своей оси (рис. 41). Более того, согласно [118], минимум потенциальной ямы сдвинут относительно центрального положения, и экситон оказывается в несимметричном положении. Такое состояние может быть описано как комбинация F -центра (электрон, захваченный вакансией; такие центры называются еще центрами окраски и обозначаются как F -центры от немецкого *Farbenzenter*) и H -центра (межузельный нейтральный ион галоида). Если энергия, необходимая для дальнейшего сдвига, достаточно мала, то такой процесс может быть первым шагом к образованию дефектов в щелочно-галоидных кристаллах. После обмена электроном между F - и H -центрами возникает два новых типа дефектов: α -центр (анионная вакансия) и I -центр (межузельный отрицательный ион аниона; такие междузельные ионы называются еще интерстициалами). Таким образом, создание автолокализованных экситонов напрямую связано с образованием дефектов под действием ультрафиолетового облучения (см., например, [113, 124, 125]). Это объясняет легкое окрашивание таких кристаллов под действием фотонов относительно малой энергии.

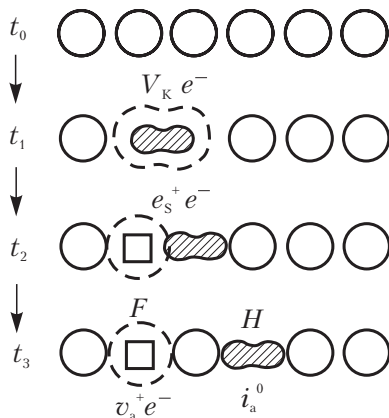


Рис. 41. Схема образования F - и H -центров из автолокализованного экситона (из [124]).

§ 15 Взаимодействие поляронов со светом

15.1 Непрямые межзонные переходы

В предыдущем параграфе было показано, что взаимодействие зонных возбуждений с деформируемой решеткой приводит к появлению нового типа возбуждений - поляронов. В случае образования сильно локализованных состояний взаимодействие с фононами наиболее сильное, что объясняется отнесенностью неподвижностью электронов и тем самым постоянным взаимодействием с поляризационной ямой. Для движущихся поляронов взаимодействие слабее — электрон (или дырка, или экситон) постоянно убегает от поляризационной ямы. Этим объясняется тот факт, что свет взаимодействует с автолокализованными возбуждениями практически так же, как и с дефектными состояниями, причем выполняется случай сильной связи. Например, люминесценция автолокализованного экситона имеет гауссову форму линии с довольно большим стоковым сдвигом. Совсем по-другому свет взаимодействует с поляронами малой массы.

Начнем исследование взаимодействия света с зонными возбуждениями со случая не прямых переходов. Как указывалось ранее, в недеформируемой решетке возможно поглощение только с сохранением импульса электрона при переходе из одной зоны в другую, поскольку волновой вектор фотона намного меньше типичного размера зоны Бриллюэна — порядка обратной постоянной решетки. В случае деформируемой решетки фононы также могут принимать участие в процессе поглощения, играя роль частицы, которой передается импульс электрона при переходе из валентной зоны в зону проводимости. Особое значение имеют такие процессы в том случае, когда порог не прямых переходов E_g^i оказывается меньше, чем для прямых E_g^d (см. рис. 42). При этом начало

фундаментального поглощения определяется непрямыми процессами.

Для простоты предположим, что электрон-фононное взаимодействие является довольно слабым, и процесс поглощения света можно рассматривать по теории возмущений как для взаимодействия с фотонами, так и для взаимодействия с фононами. Будем рассматривать переходы из валентной зоны v в зону проводимости c .

Поскольку прямые переходы между валентной зоной и зоной проводимости вблизи порога невозможны ($\hbar\omega < E_g^d$), необходимо использовать теорию возмущений второго порядка. При этом главным оказывается член, описывающий переход из зоны v в зону c через промежуточное состояние β не только с поглощением фотона, но и с поглощением или испусканием фонона. Сохранение импульса при каждом из переходов требует выполнения равенств $\mathbf{k}_v = \mathbf{k}_\beta$, $\mathbf{k}_\beta \pm \mathbf{q} = \mathbf{k}_c$, где \mathbf{k}_α — импульс электрона в каждой из рассматриваемых зон, а \mathbf{q} — импульс фотона. Знак плюс соответствует поглощению фонона, минус — испусканию. Таким образом, при переходе электрона из состояния $|v, \mathbf{k}_v\rangle$ в состояние $|c, \mathbf{k}_c\rangle$ взаимодействие происходит с испусканием или поглощением фонона из ветви s с волновым вектором $\mathbf{q} = \pm(\mathbf{k}_c - \mathbf{k}_v)$. Поэтому изменяются только числа заполнения фононов n_{qs} . Матричный элемент такого взаимодействия с поглощением (испусканием) фонона имеет вид

$$\frac{\langle c, \mathbf{k}_c; n_{qs} \mp 1 | F_s^{c\beta}(\mathbf{q}) | \beta, \mathbf{k}_v; n_{qs} \rangle \langle \beta, \mathbf{k}_v; n_{qs} | \mathbf{e} \mathbf{r} | v, \mathbf{k}_v; n_{qs} \rangle}{E_\beta(\mathbf{k}_v) - E_v(\mathbf{k}_v) - \hbar\omega}. \quad (15.1)$$

Для получения коэффициента поглощения необходимо, просуммировать по всем начальным и конечным состояниям с соответствующими множителями квадраты матричных элементов (15.1) с учетом закона сохранения энергии:

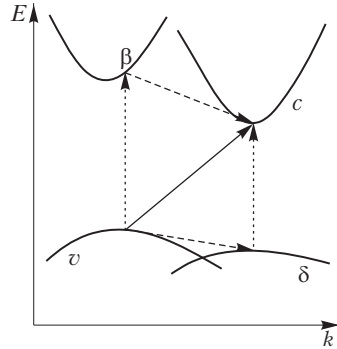


Рис.42. Схема непрямых переходов в диэлектрике.

$$\varepsilon_2(\omega) = \frac{e^2}{8\pi^4 m^2} \sum_{\beta, s, \pm} \iiint d^3 k_v d^3 k_c d^3 q \times \\ \times \delta(E_c(\mathbf{k}_c) - E_v(\mathbf{k}_v) - \hbar\omega \pm \hbar\Omega_{\mathbf{k}_s}) \delta(\mathbf{k}_v - \mathbf{k}_c \pm \mathbf{q}) \times \\ \times \frac{\left| \langle c, \mathbf{k}_v | F_s^{c\beta}(\mathbf{q}) | \beta, \mathbf{k}_v \rangle \right|^2 \left| \langle \beta, \mathbf{k}_v | \mathbf{e}_r | v, \mathbf{k}_v \rangle \right|^2 (\bar{n}_{\mathbf{q}_s} + 1/2 \pm 1/2)}{(E_\beta(\mathbf{k}_v) - E_v(\mathbf{k}_v) - \hbar\omega)^2}.$$

Здесь $F_s^{c\beta}$ описывает электрон-фононное взаимодействие, $\hbar\Omega_{\mathbf{q}_s}$ — энергия фонона ветви s с импульсом $\hbar\mathbf{q}$, $\bar{n}_{\mathbf{q}_s}$ — среднее число заполнения для фононов, зависящее от температуры: $\bar{n}_{\mathbf{q}_s} = [\exp(\hbar\Omega_{\mathbf{q}_s}/k_B T) - 1]^{-1}$. Суммирование по β введено, поскольку промежуточной зоной может быть любая, в том числе и s и v . Кроме того, в ε_2 будут вносить вклад и процессы, идущие через зону δ (см. рис. 42). Их вклад можно записать по аналогии с (15.1) с заменой $v \leftrightarrow c$ и $\beta \leftrightarrow \delta$.

Предположим, как и при выводе выражений на краю прямых разрешенных переходов, что квадрат матричного элемента (15.1) слабо зависит от \mathbf{k}_v , \mathbf{k}_c в области вблизи экстремумов. В этой области можно также положить $\mathbf{q} = \mathbf{q}_0$, где \mathbf{q}_0 — расстояние между экстремумами зон \mathbf{k}_{v0} и \mathbf{k}_{c0} в пространстве волновых векторов. Вычисления для случая изотропных зон дают

$$\varepsilon_2(\omega) \sim \sum_s C_s \left[(\hbar\omega - E_g^i + \hbar\Omega_{\mathbf{q}_0 s})^2 \theta(\hbar\omega - E_g^i + \hbar\Omega_{\mathbf{q}_0 s}) \bar{n}_{\mathbf{q}_0 s} + \right. \\ \left. + (\hbar\omega - E_g^i - \hbar\Omega_{\mathbf{q}_0 s})^2 \theta(\hbar\omega - E_g^i - \hbar\Omega_{\mathbf{q}_0 s}) (\bar{n}_{\mathbf{q}_0 s} + 1) \right], \quad (15.2)$$

где C_s — некоторые константы. Тем самым порог поглощения превращается в несколько порогов, имеющих квадратичную зависимость поглощения от энергии. Число порогов соответствует числу фононных мод, с которыми происходит взаимодействие. В силу наличия множителей $\bar{n}_{\mathbf{q}_s}$ в (15.2) коэффициент поглощения сильно зависит от температуры, причем при низких температурах идут только процессы с испусканием фононов (второй член в (15.2)). Если электронный переход между v и β (или между c и δ) в точке экстремума дипольно-запрещен, то энергетическая зависимость меняется с квадратичной на кубическую.

Как и в случае прямой запрещенной зоны, между электроном в зоне s и дыркой в зоне v может наблюдаться электростатическое взаимодействие, которое приводит к появлению непрямых экситонов. Полный импульс такого экситона близок к $\hbar\mathbf{q}_0$, а энергия (в случае экситонов большого радиуса) описывается ридберговской серией (см. § 6) с точкой сгущения E_g^i . Оптические переходы в случае разрешенных состояний происходят во все s -состояния, как и в случае прямых экситонных переходов, а острые дельтаобразные пики заменяются порогами со корневым поведением:

$$\varepsilon_2(\omega) \sim \sum_{s,\pm,n} C_s |F_n(0)|^2 (\hbar\omega - E_g^i + Ry^*/n^2 \pm \hbar\Omega_{\mathbf{q}_0s})^{1/2} \times \\ \times \theta(\hbar\omega - E_g^i + Ry^*/n^2 \pm \hbar\Omega_{\mathbf{q}_0s})(\bar{n}_{\mathbf{q}_0s} + 1/2 \mp 1/2),$$

где $|F_n(0)|^2$ — вероятность нахождения электрона и дырки в одной точке в ns -состоянии экситона.

Непрямые переходы в чаще всего связаны с поглощением и испусканием одного фонона. Однако в твердом теле наблюдается большое количество эффектов, обусловленных многофононными процессами. Одно из наиболее известных проявлений этих процессов — зависимость ширины запрещенной зоны E_g (как прямой, так и не прямой) от температуры. Известно, что с ростом температуры E_g понижается. Это связано, в частности, с тепловым расширением твердых тел и с зависимостью энергетической щели от постоянной решетки, равновесное значение которой изменяется с температурой. Однако эффект зависимости $E_g(T)$ от температуры связан также и с другими многофононными эффектами, возникающими в рамках поляронной теории.

15.2 Функция формы линии для поглощения света экситоном

Многофононными эффектами объясняется также и форма линии прямых экситонных переходов в твердом теле. Ранее в разделе 6.3 было показано, что оптические переходы в прямые разрешенные состояния приводят к острым линиям с нулевой шириной. Как и в случае взаимодействия с уровнями, обусловленными дефектами, форма линии таких переходов определяется только фононами. Однако перенесение полученных в § 13 результатов непосредственно на случай

взаимодействия экситона со светом при учете деформации решетки невозможно по той причине, что экситонные состояния образуют не уровень, а целую зону энергий. С изменением полного импульса экситона происходит и изменение энергии экситона.

Обозначим полуширину экситонной зоны через B (рис. 39). Другим параметром размерности энергии является амплитуда вызванных фононами флуктуаций потенциала. В зависимости от соотношения между B и величиной этих флуктуаций потенциала, можно различить несколько предельных случаев. Запишем общие выражения для формы линии поглощения в случае экситона в соответствии с § 13 и [67]:

$$G(\omega) \sim \int_{-\infty}^{\infty} dt \exp \left[i \left(\omega - \frac{E_0}{\hbar} t \right) - \frac{1}{\hbar^2} \int_0^t d\tau (t - \tau) g(\tau) \right], \quad (15.3)$$

где

$$g(\tau) = \sum_{\mathbf{k}} |\gamma_{\mathbf{k}}|^2 \left[\frac{e^{-i\Omega_{\mathbf{k}}\tau}}{1 - e^{-\hbar\Omega_{\mathbf{k}}/k_B T}} + \frac{e^{i\Omega_{\mathbf{k}}\tau}}{e^{\hbar\Omega_{\mathbf{k}}/k_B T} - 1} \right] e^{i(E_0 - E(-\mathbf{k}))\tau/\hbar}$$

— корреляционная функция флуктуаций потенциала из-за электрон-фононного взаимодействия. Здесь $E(\mathbf{k})$ обозначает закон дисперсии экситонов, $\gamma_{\mathbf{k}}$ — сила экситон-фононной связи; для продольных оптических фононов $\gamma_{\mathbf{k}} = 4\pi e^2 f_{\mathbf{k}}/k^2$, где $f_{\mathbf{k}}$ — формфактор экситона:

$$f_{\mathbf{k}} = \sum_{\mathbf{R}} |F(\mathbf{R})|^2 \left[\exp \left(i \frac{m_h \mathbf{kR}}{m_e + m_h} \right) - \exp \left(-i \frac{m_e \mathbf{kR}}{m_e + m_h} \right) \right], \quad (15.4)$$

$F(\mathbf{R})$ — огибающая экситонной волновой функции Ванье-Мотта, определенная в § 6. Это выражение мы записали только для поляризационного взаимодействия экситонов Ванье-Мотта с оптическими фононами, но легко получить константы $\gamma_{\mathbf{k}}$ и для любого другого типа взаимодействия экситонов с фононами.

Член с $\exp(-i\Omega_{\mathbf{k}}\tau)$ соответствует испусканию, а с $\exp(i\Omega_{\mathbf{k}}\tau)$ — поглощению оптического фонона с импульсом $\hbar\mathbf{k}$. Множитель $\exp(-iE(-\mathbf{k})\tau/\hbar)$ учитывает изменение энергии экситона при изменении импульса экситона на $-\hbar\mathbf{k}$. Выделим фигурирующий в (15.3) интеграл:

$$L(t) = \frac{1}{\hbar^2} \int_0^t d\tau (t - \tau) g(\tau) .$$

Эта функция может характеризоваться двумя временами: τ_c — временем затухания корреляционной функции $g(\tau)$, и τ_a — временем, за которое действительная часть $L(t)$ становится равной единице. Время τ_c является наименьшим из времен $2\pi/\Omega_{\mathbf{k}}$ или \hbar/B . Обычно для экситонов τ_c определяется полушириной экситонной зоны, а не частотой оптического фонона. Это означает, что перемешивание по фазам происходит за счет дисперсии экситона. Введем также амплитуду флуктуирующих воздействий: $D^2 = g(0)$. Для малых времен $t \ll \tau_c$ можно получить

$$L(t) \approx D^2 t^2 / 2\hbar^2 , \quad (15.5)$$

а для больших $|t| \gg \tau_c$ —

$$L(t) = i\Delta_0 t / \hbar - \Gamma_0 |t| / \hbar ,$$

где $\Gamma_0 \approx D^2 \tau_c / \hbar$ и Δ_0 — ширина линии и энергетический сдвиг за счет взаимодействия с фононами.

Случай $\tau_c \gg \tau_a$ ($\tau_a \approx \hbar/D$ из (15.5)) называется пределом медленной модуляции, и линия поглощения имеет гауссову форму:

$$\begin{aligned} G(\omega) &\sim \int_{-\infty}^{\infty} dt \exp \left[i \left(\omega - \frac{E_0}{\hbar} t \right) - \frac{D^2}{2\hbar^2} t^2 \right] \sim \\ &\sim \frac{1}{\sqrt{2\pi D}} \exp \left[- \frac{(\hbar\omega - E_0)^2}{2D^2} \right] . \end{aligned}$$

В обратном предельном случае, $\tau_c \ll \tau_a$, $g(\tau)$ отлично от нуля только в самом начале области интегрирования в $L(t)$, и этот случай называется пределом быстрой модуляции. При этом форма линии имеет лоренцеву форму:

$$G(\omega) \sim \frac{1}{\pi} \frac{\Gamma_0}{(\hbar\omega - E_0 - \Delta_0)^2 + \Gamma_0^2} .$$

В случае $D \gg \hbar\Omega$ эти два предельных случая соответствуют $D \gg B$ (гауссова линия) и $D \ll B$ (лоренцова линия). В

первом случае форма линии соответствует взаимодействию локализованного экситона с фононами (ср. с пекаровскими линиями поглощения), во втором возникает уширение уровня экситона с $\mathbf{k} = 0$ за счет рассеяния на фононах. Ширина лоренцевой линии Γ_0 много меньше D , поэтому гауссова линия много шире лоренцевой.

Описанные выше два предела быстрой и медленной модуляции типичны для всех случаев взаимодействия системы с флуктуирующими полями и проявляются во всех областях физики. Естественно, что реальная ситуация не описывается ни одним из этих пределов. Эти формы линии справедливы только достаточно близко от центра линии поглощения. Хвосты линий поглощения никогда не соответствуют этим двум пределам. Поглощение затухает быстрее, чем описываемое лоренцевой линией, но медленнее описываемой гауссовой. Форма линии в ряде случаев оказывается асимметричной.

Полуширина экситонной зоны B играет существенную роль и в процессе локализации экситона в решетке (см., например, [113]). Как видно из рис. 39, при изменении конфигурационной координаты воспринимающей моды Q_1 имеется минимум, смещенный из начала координат. Выигрыш энергии при релаксации решетки вокруг экситона обозначим как E_{LR} . Однако при локализации экситон должен перейти в состояние, в котором его импульс не является определенным. При этом энергия экситона повышается на величину B по сравнению с минимумом. В том случае, когда $E_{LR} < B$, автолокализованное состояние экситона является не стабильным, а лишь метастабильным. В обратном случае $E_{LR} > B$ автолокализованное состояние является стабильным, а состояние свободного экситона — метастабильным. Эти состояния разделены между собой потенциальным барьером, поэтому в принципе возможно и сосуществование двух типов экситонов. При исследовании свечения таких экситонов видно, что они существенно разделяются по положению и полуширине линии люминесценции. Автолокализованные экситоны характеризуются большим стоковым сдвигом (порядка $2E_{LR} - B$) и значительной шириной линии, в то время как свечение свободных экситонов имеет узкую резонансную линию вблизи экситонной линии поглощения.

15.3 Правило Урбаха

Низкоэнергетичный хвост фундаментального поглощения часто описывается эмпирическим правилом Урбаха [94], согласно которому

$$\varepsilon_2(\omega) \sim \exp \frac{\sigma(\hbar\omega - E_{ex})}{k_B T} \text{ для } \hbar\omega < E_{ex}. \quad (15.6)$$

Величина σ для разных кристаллов принимает значения от 0,5 до 2, E_{ex} — энергия, примерно соответствующая максимуму экситонного поглощения. Правило Урбаха выполняется для большого числа веществ в диапазоне коэффициентов поглощения от 10 до 10^5 см^{-1} . Имеется много работ, в которых выводится соотношение (15.6) на основании различных исходных посылок. По-видимому, в идеальных кристаллах это соотношение является результатом многофононных процессов, в частности, временных флуктуаций индуцированного фононами потенциала, в то время как в кристаллах с примесями и неупорядоченных системах причина применимости этого правила связана с пространственными флуктуациями потенциала. Если этот эффект связан с многофононным взаимодействием, температура в знаменателе экспоненты в (15.6) заменяется на типичную для взаимодействия с фононами величину:

$$\varepsilon_2(\omega) \sim \exp \frac{\sigma(\hbar\omega - E_{ex})}{\hbar\Omega \text{cth}(\hbar\Omega/2k_B T)} \text{ для } \hbar\omega < E_{ex}.$$

Здесь Ω — средняя частота фононов, взаимодействующих с экситонами.

Литература

1. Давыдов А. С., Теория твердого тела. М.: Наука, 1976.
2. Бассани Ф., Пастори Паррavicchini Дж., Электронные состояния и оптические переходы в твердых телах. М.: Наука, 1982.
3. Нокс Р., Теория экситонов. М.: Мир, 1966.
4. Ансельм А. И., Введение в теорию полупроводников. М.: Наука, 1978.
5. Стоунхэм А. М., Теория дефектов в твердых телах. Т. 1, 2. М.: Мир, 1978.
6. Бонч-Бруевич В. Л., Калашников С. Г., Физика полупроводников. М.: Наука, 1977.
7. Займан Дж., Принципы теории твердого тела. М.: Мир, 1974.
8. Займан Дж., Вычисление блоховских функций. М.: Мир, 1978.
9. Фэн Г., Фотон-электронное взаимодействие в кристаллах в отсутствие внешних полей. М.: Мир, 1969.
10. Агранович В. М., Гинзбург В. Л., Кристаллооптика с учетом пространственной дисперсии и теория экситонов. 2-е изд. М.: Наука, 1979.
11. Агранович В. М., Галанин М. Д., Перенос энергии электронного возбуждения в конденсированных средах. М.: Наука, 1978.
12. Синхротронное излучение, свойства и применения. Под ред. К. Кунца. М.: Мир, 1981.
13. Фок М. В., Введение в кинетику люминесценции кристаллофосфоров. М.: Наука, 1964.
14. Антонов-Романовский В. В., Кинетика фотолуминесценции кристаллофосфоров. М.: Наука, 1966.
15. Кюри Д., Люминесценция кристаллов. М.: ИЛ, 1961.
16. Филипс Дж., Оптические спектры твердых тел в области собственного поглощения. М.: Мир, 1968.
17. Алукер Э. Д., Лусис Д. Ю., Чернов С. А., Электронные возбуждения и радиолуминесценция щелочно-галогидных кристаллов. Рига: Зинатне, 1979.
18. Васильев А.Н., Михайлин В.В., Введение в спектроскопию твердого тела, М., из-во Московского университета, 1987
19. Ландау Л.Д., Лифшиц Е.М., Теоретическая физика, т. II, Теория поля, М., Наука, 1973
20. Ландау Л.Д., Лифшиц Е.М., Теоретическая физика, т. VIII, Электродинамика сплошных сред, М., Наука, 1982.
21. Ландау Л.Д., Лифшиц Е.М., Теоретическая физика, т. V, Статистическая физика, часть I, М., Наука, 1976
22. Климонтович Ю.Л., Кинетическая теория неидеальных газов и неидеальной плазмы, М., Наука, 1982
23. Борн М., Вольф Э., Основы оптики. М.: Наука, 1969.
24. Tamm I.E., Zs. Physik v. 76, p. 849, 1932.
25. Davison S.G. and Levine J.D., Surface States, New York and London: Academic Press, 1970.
26. Kramers H.A. Atti Congr. Inten. Fis. Como. v. 2, p. 545, 1927.
27. Kronig R.J. Opt. Soc. Am., v. 12, p. 547, 1926

28. Stern F. 1963 in Solid State Physics 15 eds Seitz F and Tumbull D (New York London: Academic Press)
29. Roessler D.M. Br. J. Appl. Phys. v. 16, p. 1119, 1965.
30. Mott N.F. and Jones H., Theory of the Properties of Metals and Alloys (Oxford) 1956.
31. Bortz M.L. and French R.H. Applied Spectroscopy, Nov.-Dec., 1989
32. Васильев А.Н., Козлов А.В., Степанюк В.С. и Качнельсон А.А., Вестник Московского университета, сер. 3, Физика, Астрономия, т. 31, стр. 76–81, 1990.
33. Корн Г.А. и Корн Т.М., Справочник по математике для студентов и инженеров, A and Korn T M 1968 Mathematical Handbook for Students and Engineers (New York McGraw-Hill)
34. Васильев А. Н., Колобанов В. Н., Куусманн И. Л., Лущик Ч. Б., Михайлин В. В., ФТТ, т. 27, с. 2696, 1985.
35. Altarelli M., Dexter D.L., Nussenzveig H.M. and Smith D.Y. Phys. Rev. v. B6, p. 4502, 1972.
36. Shpinkov I.N., Kamenskikh I.A., Kirm M., Kolobanov V.N., Mikhailin V.V., Vasil'ev A.N. and Zimmerer G., Phys. Stat. Sol. (a), v. 170, p. 167, 1998.
37. Neufeld J. and Ritchie R.H., Phys. Rev., v. 98, p. 1632, 1965.
38. Vager Z. and Gemmel D.S., Phys. Rev. Lett. v. 37, p. 1352, 1976.
39. Калашников Н.П., Ремизович В.С. и Рязанов М.И., Столкновения быстрых заряженных частиц в твердых телах, Москва, Наука, 1980.
40. Philipp H.R. and Ehrenrich H., Phys. Rev., v. 129, p. 1550, 1963.
41. Ritchie R.H., Phys. Rev., v. 106, p. 874, 1957.
42. Пайнес Д., Элементарные возбуждения в твердых телах, Москва, Мир, 1965.
43. Pines D. and Nozieres P., The Theory of Quantum Liquids, New York: Benjamin, 1966.
44. Platzman P.M. and Wolff P.A., Waves and Interactions in Solid State Plasmas, New York: Academic Press, 1973.
45. Ю П., Кардона М., Основы физики полупроводников, Москва, Физматлит, 2002.
46. Киржниц Д.А., УФН, т. 119, с. 367, 1976.
47. Долгов О.В. и Максимов Е.С., УФН, т. 135, с. 441, 1981.
48. Климонтович Ю.Л., Кинетическая теория электромагнитных процессов, М., Наука, 1980.
49. Ландау Л.Д., Лифшиц Е.М., Теоретическая физика, т. III, Квантовая механика, Нерелятивистская теория, М.: Наука, 1974
50. Nakajima S. Progr. Theor. Phys. (Kyoto), v. 20, p. 948, 1958.
51. Zwanzig R., in Lectures Theoretical Physics (Boulder), v. 3, p. 160, New York: Interscience Publishers, 1961.
52. Grabert H., Projection Operator Techniques in Nonequilibrium Statistical Mechanics, Springer Tracts in Modern Physics, v. 95, Springer-Verlag, 1982.
53. Силин В.П. и Рухадзе А.А., Электромагнитные процессы в плазме и плазмоподобных средах, М., Атомиздат, 1961.
54. Лифшиц Е.М., Питаевский Л.П., Теоретическая физика, т. X, Физическая кинетика, М., Наука, 1979
55. Лифшиц Е.М., Питаевский Л.П., Теоретическая физика, т. IX, Статистическая физика, т. 2, Теория конденсированного состояния, М., Наука, 1978
56. Callaway J., Energy Band Theory, New York, 1964.
57. Маделунг О., Физика твердого тела, т. 1, 2, М.: Наука, 1985.

58. Слэтер Дж., Методы самосогласованного поля для молекул и твердых тел, М.: Мир, 1978.
59. Хаммермеш М., Теория групп и ее применение к физическим проблемам, М.: Мир, 1966.
60. Bouckaert L.P., Smoluchowski R. and Wigner E.P., Phys. Rev., v. 50, p. 58, 1936.
61. Van Hove L., Phys. Rev., v. 89, p. 1189, 1953.
62. Pantelides S.T., Phys. Rev. B, v. 11, p. 5082, 1975.
63. Mott N.F. and Davis E.A., Electronic Processes in Non-Crystalline Materials, Oxford: Clarendon Press, 1971.
64. Brust D., Phys. Rev., v. 134, p. A1337, 1964.
65. Степанюк В.С., Григоренко А.А., Козлов А.В., Фарберович О.В., Михайлин В.В., Степанова Е.В., ФТТ, т. 31, с. 58, 1989.
66. Амусья М.Я., Атомный фотоэффект, М.: Наука, 1987.
67. Котани А., Тойозава И., в сборнике Синхротронное излучение. Свойства и применение, под ред. К.Кунца, с. 211, М.: Мир, 1981.
68. Колдинг К., в сборнике Синхротронное излучение. Свойства и применение, под ред. К.Кунца, с. 278, М.: Мир, 1981.
69. Sayers D.E., Lytle F.W. and Stern E.A., in Advances in X-Ray Analysis, ed. by Henke B.L., Newkirk J.B. and Mallett G.R., New York: Plenum Press, p. 248, 1970.
70. Wannier G.H., Phys. Rev., v. 52, p. 191, 1937.
71. Нокс Р., Теория экситонов. М.: Мир, 1966.
72. Frenkel J., Phys. Rev., v. 37, p. 17, 1931.
73. Heller W.R. and Marcus A., Phys. Rev., v. 84, p. 809, 1951.
74. Давыдов А.С., ЖЭТФ, т. 18, с. 210, 1958.
75. Elliott R.J., Phys. Rev., v. 108, p. 1384, 1957.
76. Macfarlane G.G., McLean T.R., Quarrington J. E. and Roberts V. Phys. Rev., v. 111, p. 1245, 1958.
77. Hayashi M. and Katsuki K., Journ. Phys. Soc. Japan, v. 5, p. 381, 1950.
78. Dresselhaus G.J., Phys. Chem. Solids, v. 1, p. 14, 1956.
79. Haken H. and Schottky W., Zs. Phys. Chem., v. 16, p. 218, 1958.
80. Toyozawa Y., Progr. Theoret. Phys. (Kyoto), v. 12, p. 421, 1954.
81. Hilsh R. and Pohl R. W., Zs. Phys., v. 48, p. 384, 1928.
82. Vambynek W. et al, Rev. Mod. Phys., v. 44, p. 716, 1972.
83. Landsberg P.T. and Adams M.J., Journ. Luminesc., v. 7, p. 3, 1973.
84. Fano U., Phys. Rev., v. 124, p. 1866, 1961.
85. Shibatani A. and Toyozawa Y., J. Phys. Soc. Japan, v. 25, p. 335, 1968.
86. Haensel R., Keitel G., Kunz C. and Schreiber P., Phys. Rev. Lett., v. 25, p. 208, 1970.
87. Kittel C. and Mitchel A.H., Phys. Rev., v. 96, p. 1488, 1954.
88. Ham F.S., Sol. St. Phys. II, 1955.
89. Seaton M.J., Mon. Not. R. Astr. Soc., v. 118, p. 504, 1958.
90. Bebb H.B., Phys. Rev., v. B5, p. 4201, 1972.
91. Bethe H.A. and Salpeter E.E., Quantum Mechanics of One- and Two-Electron Atoms, New York, 1957.
92. Займан Дж., Модели беспорядка, М.: Мир, 1982.
93. Лифшиц И.М., Гредескул С.А., Пастур Л.А., Введение в теорию неупорядоченных систем, М.: Наука, 1982
94. Urbach F., Phys. Rev., v. 92, p. 1324, 1953.
95. Пекар А.М., ЖЭТФ, т. 20, с. 510, 1950.
96. Huang K. and Rhys A., Proc. Roy. Soc., v. A204, p. 406, 1950.

97. Jahn H.A. and Teller E., Proc. Roy. Soc., v. A161, p. 220, 1937.
98. Condon E.U., Phys. Rev., v. 32, p. 358, 1928.
99. Frank J., Trans. Faraday Soc., v. 21, p. 536, 1925.
100. Braisford A.D. and Chang T.Y., J. Chem. Phys., v. 53, p. 3108, 1970.
101. Freed K.F. and Jortner J., J. Chem. Phys., v. 52, p. 6272, 1970.
102. Riseberg L.A. and Moos H.W., Phys. Rev., v. 174, p. 429, 1968.
103. Shockley W. and Bardeen J., Phys. Rev., v. 77, p. 407, 1949.
104. Frohlich H. Adv. in Phys., v. 3, p. 325, 1954.
105. Frohlich H., in Polarons and Excitons (eds. Koper C.G. and Whitfield G.D.), Edinburgh: Oliver and Boyd, 1962.
106. Feynman R.P., Phys. Rev., v. 97, p. 660, 1955.
107. Feynman R.P., Statistical Mechanics, Massachusetts: Benjamin, 1972.
108. Landau L.D., Phys. Z. Sowjetunion, v. 3, p. 664, 1933.
109. Пекар С.И., Исследования по электронной теории кристаллов, М.: Гостехиздат, 1951.
110. Toyozawa Y., Progr. Theor. Phys., v. 26, p. 29, 1961.
111. Toyozawa Y., Exciton Lattice Interaction. Fluctuation, Relaxation and Defect Formation. Technical Report ISSP, Japan. Ser. a, N648, 1974.
112. Toyozawa Y., in Relaxation of Elementary Excitations, ed. by Kubo R. and Hanamura E., Berlin-Heidelberg-New York:Springer, 1980.
113. Toyozawa Y., Semiconductors and Insulators, v. 5, p. 175, 1983.
114. Sumi A., J. Phys. Soc. Japan, v. 43, p. 1286, 1977.
115. Рашба Е.И., Оптика и спектроскопия, т. 2, с. 75, 1957.
116. Рашба Е.И., Изв. АН СССР, сер. физ., т. 40, с. 1973, 1976.
117. Mott N.F. and Stoneham A.M., J. Phys. C: Solid State Phys., v. 10, p. 3391, 1977.
118. Song K.S., Williams R.T., Self-Trapped Excitons, Springer, 1995.
119. Kanzig W., Phys. Rev., v. 99, p. 1890, 1955.
120. Castner T.G. and Kanzig W., J. Phys. Chem. Sol., v. 3, p. 178, 1957.
121. Seidel H. and Wolf H.C., in Physics of Color Centers, New York-London: Academic Press, p. 537, 1968.
122. Kabler M.N., in Point Defects in Solids, New York:Plenum Press, p. 327, 1972
123. Алукер Э.Д., Лусис Д.Ю., Чернов С.А., Электронные возбуждения и радиолуминесценция щелочно-галогидных кристаллов. Рига: Зинатне 1979
124. Лущик Ч.Б. и Лущик А.Ч., Распад электронных возбуждений с образованием дефектов в твердых телах, М.:Наука, 1989.
125. Itoh N., Semiconductors and Insulators, v. 5, p. 165, 1983.

Оглавление

Глава 1. Взаимодействие электромагнитного поля с твердыми телами	7
§ 1. Феноменологическое описание взаимодействия света с веществом	7
1.1 Оптические функции твердых тел	7
1.2 Отражение света от твердых тел	18
1.3 Соотношения Крамерса – Кронига и правила сумм	25
§ 2. Характеристические потери энергии электронов	34
§ 3. Микроскопическое описание взаимодействия света с веществом	38
3.1 Диэлектрическая проницаемость	38
3.2 Поляризационное приближение для взаимодействия электромагнитных полей с твердым телом	45
3.3 Флуктуационно-диссипационная теорема	53
3.4 Кинетическое уравнение в поляризационном приближении	56
Глава 2. Состояния в недеформируемой решетке	61
§ 4. Электронные состояния в недеформируемой решетке	61
4.1. Иерархия приближений в теории твердого тела	61
4.2 Одноэлектронное приближение	62
4.3 Приближение почти свободных электронов	69
4.4 Приближение сильно связанных электронов	71
4.5 Плотность зонных состояний в твердом теле	77
4.6 Заполнение одноэлектронных состояний в твердом теле	79
§ 5. Переходы под действием света в идеальном кристалле	81
5.1 Межзонные переходы	81
5.2 Пространственная дисперсия межзонной диэлектрической проницаемости	89
5.3 Эффект Франца-Келдыша	91
5.4 Поглощение фотонов большой энергии	92
5.5 EXAFS	95
§ 6. Экситонные эффекты в твердых телах	99
6.1 Кулоновское взаимодействие электрона и дырки	99
6.2 Экситон Френкеля	104
6.3 Экситон Ванье-Мотта	108
6.4 Экситоны в реальных кристаллах	117

§ 7.	Конфигурационное взаимодействие	123
7.1	Диэлектрическая проницаемость системы взаимодействующих частиц	123
7.2	Эффект Оже	125
7.3	Эффект Фано	130
§ 8.	Поляритоны	135
§ 9.	Электронные состояния дефектов в недеформируемой решетке	141
9.1	Классификация дефектов в твердых телах	141
9.2	Дефекты большого радиуса	142
9.3	Глубокие дефекты	145
9.4	Плотность состояний в кристалле с дефектами	149
§ 10.	Фотоионизация дефектов	152
§ 11.	Неупорядоченные системы	157
Глава 3. Электронные состояния в деформируемой решетке		162
§ 12.	Фононы в кристаллах	162
12.1	Фононы в идеальном кристалле	162
12.2	Фононы в кристаллах с дефектами	169
§ 13.	Переходы между состояниями дефектов при учете фононов	173
13.1	Функция формы линии поглощения при учете фононов	173
13.2	Слабая электрон-фононная связь	178
13.3	Сильная электрон-фононная связь	180
13.4	Безызлучательные переходы	185
§ 14.	Поляроны	187
14.1	Поляроны большого радиуса. Деформационное электрон-фононное взаимодействие	187
14.2	Поляроны большого радиуса. Поляризационное взаимодействие	191
14.3	Континуальное приближение	194
14.4	Автолокализованные дырки и экситоны в щелочно-галогидных кристаллах. Образование дефектов	201
§ 15.	Взаимодействие поляронов со светом	204
15.1	Непрямые межзонные переходы	204
15.2	Функция формы линии для поглощения света экситоном	207
15.3	Правило Урбаха	211
	Литература	212

Учебное пособие

Андрей Николаевич Васильев
Виталий Васильевич Михайлин

Введение в спектроскопию диэлектриков, часть I

Работа поступила в ОНТИ 30.10.2008

Тираж экз.