

2. Эффект Келдыша — Франца [20, 21]

Электрическое поле \mathcal{E} вызывает наклон краев зон. Электрон, движущийся внутри зоны от ее края, приобретает на расстоянии x кинетическую энергию $q\mathcal{E}x$ (при фиксированной полной энергии). Он обладает действительным волновым вектором и его волновая функция похожа на функции, которыми пользуются для описания зон. Если бы электрон мог двигаться от края разрешенной внутрь запрещенной зоны, то (при постоянной полной энергии) его кинетическая энергия стала бы отрицательной, а соответствующий волновой вектор — мнимым. Волновая функция электрона в запрещенной зоне e^{ikx} (здесь k — мнимое число) описывает затухающую волну. Следовательно, вероятность найти электрон в пределах запрещенной зоны уменьшается по закону:

$$\exp\left(-\frac{|E-E_e|}{\Delta E}\right),$$

где E — энергия электрона в точке x ; E_e — энергия, соответствующая краю зоны в точке x , а ΔE — параметр, зависящий от величины поля. Поэтому состояния у краев зоны уширены, это уширение оказывается экспоненциальным и характеризуется средним смещением края ΔE , зависящим от поля. Можно показать, что зависимость имеет вид [22]

$$\Delta E = \frac{3}{2} (m^*)^{-1/3} (q\hbar\mathcal{E})^{2/3}. \quad (2.4)$$

В дальнейшем мы покажем, насколько важен этот эффект для понимания оптических свойств и явлений переноса в полупроводниках (и изоляторах).

деленные от поверхности 10-микронным слоем майлара [15]. Противоположная сторона образца делается шероховатой, чтобы уменьшить отражение от задней поверхности. Эта методика позволяет регистрировать весьма малые относительные изменения отражения ($\Delta R/R \approx 5 \cdot 10^{-8}$).

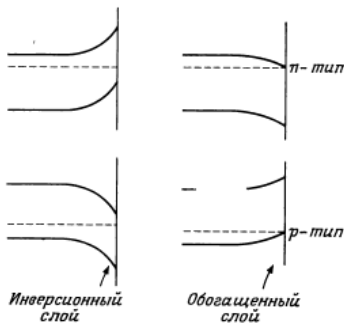
При интерпретации данных по электроотражению следует проявлять большую осторожность. Если поле приложено перпендикулярно поверхности, то оно неоднородно внутри полупроводника. Эта неоднородность сильно влияет на форму спектра за счет эффектов усреднения и смещения действительной и мнимой частей диэлектрической функции [16]. На фиг. 18.4 показано сильное влияние неоднородности электрического поля на вид осциллирующей части диэлектрической функции (две верхние кривые относятся к случаю однородного поля). Экспериментальное доказательство влияния неоднородного поля на спектр электроотражения подтвердило необходимость учета неоднородности поля при интерпретации данных [17].

Поскольку при синхронном детектировании неизбежно происходит некоторое усреднение, то существенно, чтобы модуляция электрического поля осуществлялась при помощи волн прямоугольной формы с изменением поля между двумя фиксированными значениями, так чтобы спектр, полученный при помощи такой дифференциальной методики, соответствовал определенному значению поля (желательно однородного).

Электрическое поле можно приложить к поверхности более простым способом, так чтобы спектр, полученный при помощи такой дифференциальной методики, соответствовал определенному значению поля (желательно однородного).

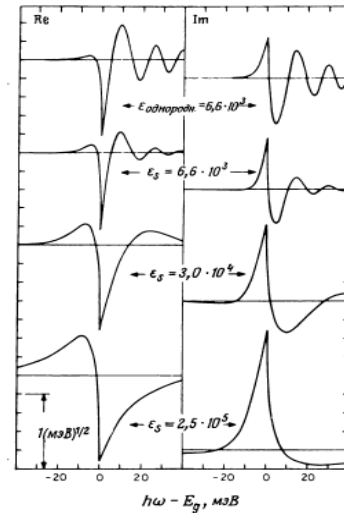
Электрическое поле можно приложить к поверхности более простым способом, так чтобы спектр, полученный при помощи такой дифференциальной методики, соответствовал определенному значению поля (желательно однородного).

Фиг. 18.5. Изменение внутреннего электрического поля на различных поверхностях. Величина электрического поля определяется наклоном краев зон.



1. Электроотражение

Электрическое поле может быть приложено вдоль поверхности полупроводника. Для этой цели между двумя электродами, напыленными на переднюю поверхность, прикладывается переменное напряжение, так чтобы электрическое поле существовало в области между параллельными краями электродов [14]. Если толщина материала мала по сравнению с расстоянием между электродами, то поле в кристалле приблизительно однородно. В этом случае данные, характеризующие отражение, можно сравнить с данными по пропусканию, полученными в тех же условиях. Отметим, что при такой конфигурации можно изменять направление поляризации падающего излучения по отношению к направлению модулирующего электрического поля.



Обычно, однако, электрическое поле направлено перпендикулярно поверхности, т. е. почти параллельно направлению распространения света. Например, для того чтобы приложить к образцу переменное электрическое поле, можно использовать прозрачные электроды, от-

Фиг. 18.4. Действительная и мнимая части индцированных полем изменений диэлектрической функции германия в области прямых межзональных переходов для различных полях.

Верхняя пара кривых соответствует приближенно однородному полю. Величина поля дана в В/см; $\langle \Delta \epsilon_1 \rangle$, отложенная по левой оси ординат, пропорциональна $\Delta E/E$, а $\langle \Delta \epsilon_2 \rangle \approx (\pi c/\omega) \langle \Delta \alpha \rangle$ (отложена по правой оси ординат [16]).

бом, а именно путем погружения полупроводника в электролит, например слабый раствор KCl в воде [18]. Переменное напряжение прикладывается между образцом и платиновым электродом, погруженным в электролит. Излучение попадает на образец через окно в ячейке. Коротковолновая граница пропускания воды составляет 1700 Å; в длинноволновой области спектра можно работать вплоть до длин волн 2,6 мкм в системе с тонким слоем водного раствора или с органическим растворителем, не содержащим радикалов ОН [19]. Эта простая техника позволяет производить скальвание образца в электролите, не подвергая его воздействию воздуха. К образцу можно приложить механические напряжения для исследования влияния деформации на зонную структуру [20]. Применение электролитов ограничено, однако, областью температур выше 200 К.

Отметим, что поверхностный потенциал, а следовательно, и электрическое поле на поверхности и вблизи нее зависят от поверхностных состояний и природы адсорбированных примесей, как это показано на фиг. 18.5. Внешнее поле индуцирует поверхностный заряд, который может усилить или ослабить изгиб зон или даже изменить полярность поверхностного поля. Поскольку изгиб зон вблизи поверхности полупроводников n - и p -типа может иметь разные знаки, то пики в дифференциальном спектре отражения могут иметь противоположные полярности в материалах с разным типом проводимости.

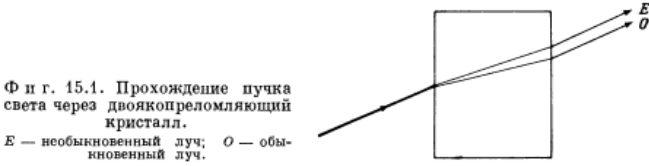
Идеальная ситуация, когда измерения производятся в условиях плоских зон, может быть получена, если к полупроводнику приложить постоянное напряжение, подобранное таким образом, чтобы поверхностный потенциал был равен нулю. В том случае, если зоны вблизи поверхности «выпрямлены», сигнал электроотражения проходит через минимум [11].

Поляризаційні явища

В кубічних напівпровідниках випромінювання розповсюджується в будь-якому напрямку з однаковою швидкістю c/n . В некубічних напівпровідниках оптичні властивості не ізотропні; діелектрична проникність представляє собою тензор, елементи якого залежать від напрямку. Оптична анізотропія особливо сильно впливає на поляризаційні характеристики розповсюджується хвилі. Одноосне напруження, електричне або магнітне поле можуть індукувати оптичну анізотропію в кристалі, являючись в звичайних умовах ізотропним. Нижче ми даємо короткий огляд різних оптичних властивостей неізотропних речовин, для того щоб пояснити терміни, використовувані при розгляді поляризаційних ефектів, а також вкажемо, яким чином ці ефекти можна індукувати або модулювати за допомогою зовнішніх впливів.

§ 1. ДВОЙНЕ ЛУЧЕПРЕЛОМЛЕННЯ

Явище подвійного лучепреломлення складається в тому, що хвилі з різною поляризацією розповсюджуються в речовині з різними швидкостями. В результаті світловий пучок, падаючий на



параллелепипед под произвольным углом, разделяется, давая на выходе два параллельных пучка (фиг. 15.1). Эти два пучка имеют различную поляризацию, и каждый из них характеризуется своим показателем преломления, входящим в выражение для закона Снеллиуса. При этом один из этих лучей, обыкновенный, в противоположность другому, необыкновенному, характери-

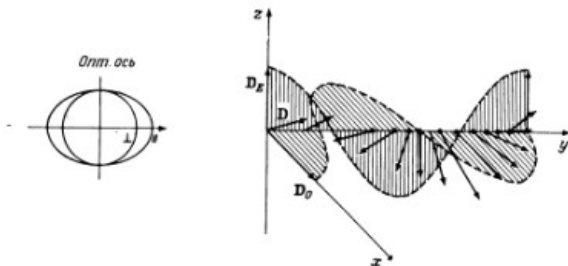
видно, что фазы двух лучей совпадают лишь при распространении вдоль оптической оси, когда оба показателя преломления одинаковы.

Плоскость, проходящая через оптическую ось и обыкновенный луч, называется главной плоскостью обыкновенного луча. Плоскость, содержащая оптическую ось и необыкновенный луч, называется главной плоскостью необыкновенного луча. Эти две плоскости пересекаются на оптической оси и в общем случае не совпадают. Необыкновенный луч поляризован в главной плоскости необыкновенного луча, тогда как обыкновенный луч поляризован перпендикулярно своей главной плоскости. Следовательно, обыкновенный луч всегда поляризован перпендикулярно оптической оси.

В качестве примера одноосного двоякопреломляющего полупроводника можно привести селен, у которого $n_o = 2,78$ и $n_e = 3,58$ в области между 9 и 23 мкм [1].

2. Эллиптическая поляризация

Пусть свет распространяется вдоль направления, перпендикулярного оптической оси одноосного кристалла. Как видно из



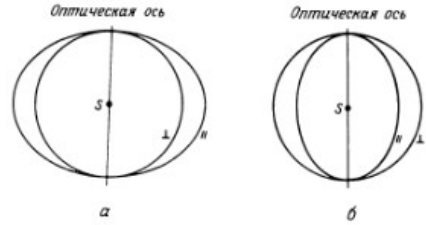
фиг. 15.3, скорости распространения обыкновенного и необыкновенного лучей различны. Следовательно, между двумя лучами будет возникать разность фаз. После того как волны прошли расстояние d , эта разность фаз равна

$$\Delta\phi = \frac{2\pi}{\lambda} d(n_o - n_e). \quad (15.1)$$

зуется показателем преломления, не зависящим от угла падения. Если свет падает на двоякопреломляющий кристалл под прямым углом, то обыкновенный луч проходит через кристалл, не меняя направления и не смещаясь, тогда как необыкновенный луч выходит из кристалла параллельно обыкновенному, но оказывается, вообще говоря, смещенным.

1. Двойное лучепреломление в одноосных кристаллах

В одноосных кристаллах показатели преломления обыкновенного и необыкновенного лучей одинаковы лишь при распространении света вдоль одного направления, называемого «оптической осью». Таким образом, оба луча распространяются с одинаковой



скоростью вдоль оптической оси, тогда как при распространении под углом к оси их скорости отличаются. Для отрицательных одноосных кристаллов $n_e < n_o$; для положительных $n_o < n_e$ (n_e и n_o — показатели преломления соответственно для необыкновенного и обыкновенного лучей). Эти свойства иллюстрируются на фиг. 15.2, где круг представляет собой поверхность равной фазы обыкновенной сферической волны, распространяющейся из центра, а овал — необыкновенной волны, распространяющейся из того же центра S. Эти две волны характеризуются взаимно перпендикулярными направлениями поляризации (направлением вектора смещения D); для обыкновенного луча проекция вектора D на направление оптической оси всегда равна нулю.

Обыкновенный луч, подчиняющийся закону Снеллиуса, всегда находится в плоскости, образованной падающим лучом и нормалью к поверхности кристалла в точке падения. Необыкновенный луч лежит, вообще говоря, не в этой плоскости. На фиг. 15.2

скоростью вдоль оптической оси, тогда как при распространении под углом к оси их скорости отличаются. Для отрицательных одноосных кристаллов $n_e < n_o$; для положительных $n_o < n_e$ (n_e и n_o — показатели преломления соответственно для необыкновенного и обыкновенного лучей). Эти свойства иллюстрируются на фиг. 15.2, где круг представляет собой поверхность равной фазы обыкновенной сферической волны, распространяющейся из центра, а овал — необыкновенной волны, распространяющейся из того же центра S. Эти две волны характеризуются взаимно перпендикулярными направлениями поляризации (направлением вектора смещения D); для обыкновенного луча проекция вектора D на направление оптической оси всегда равна нулю.

Обыкновенный луч, подчиняющийся закону Снеллиуса, всегда находится в плоскости, образованной падающим лучом и нормалью к поверхности кристалла в точке падения. Необыкновенный луч лежит, вообще говоря, не в этой плоскости. На фиг. 15.2

Если направление поляризации падающего света составляет 45° с оптической осью, то волну удобно разложить на обыкновенную составляющую с вектором D_o , перпендикулярным к оптической оси, и необыкновенную составляющую, вектор D_e которой параллелен оптической оси. Оба вектора осциллируют с оптической частотой. Поскольку скорости распространения D_e и D_o различны, их сумма (вектор D) вращается вокруг направления распространения. В данном случае (падающее излучение линейно поляризовано под углом 45° к оптической оси) конец вектора D описывает, вообще говоря, эллипс в плоскости, через которую свет выходит из кристалла. Возможны, однако, два предельных случая эллиптической поляризации, когда соответствующая разность фаз между двумя лучами приводит либо к линейной, либо к круговой поляризации.

Если толщина кристалла такова, что разность хода между обыкновенным и необыкновенным лучами составляет целое число половин, то выходящий из кристалла свет будет также линейно поляризован. Если это число нечетное, то направление поляризации выходящего света перпендикулярно направлению поляризации падающего.

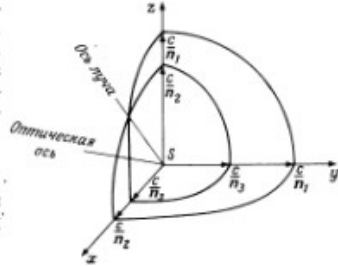
Когда разность хода равна нечетному числу четвертей длины волны, выходящий свет поляризован по кругу. При увеличении нечетного числа четвертей длины волны получаются попеременно правая и левая круговые поляризации. Пластина в четверть волны представляет собой двоякопреломляющий кристалл, толщина которого подобрана таким образом, чтобы обеспечить на нужной длине волны разность фаз $\pi/4$ между обыкновенным и необыкновенным лучами. Такая пластина превращает линейно поляризованный свет в циркулярно поляризованный и наоборот.

3. Двойное лучепреломление в двuosных кристаллах

В двuosных кристаллах имеются две оптические оси, вдоль которых один квазиобыкновенный и необыкновенный лучи распространяются с одинаковой скоростью и, следовательно, характеризуются одним и тем же значением n . Как будет ясно из дальнейшего, в двuosных кристаллах падающий луч всегда расщепляется на два луча, причем оба являются необыкновенными — за исключением того случая, когда в каждой из трех координатных плоскостей (x, y; y, z; z, x) один из двух лучей обладает свойствами обыкновенного луча; мы назовем его квазиобыкновенным лучом, поскольку как таковой он может существовать лишь в трех определенных плоскостях (фиг. 15.4).

«Лучевая ось» есть направление, которому соответствует только одна скорость распространения. Таких осей — две. Опти-

ческая ось есть направление, вдоль которого нормали к фронту волны распространяются с одинаковой скоростью. Нормаль к волновому фронту можно найти, проведя касательную к двум волновым поверхностям. Радиус-вектор, проведенный в точку касания с фронтом квазиобыкновенной волны, представляет собой оптическую ось. Оптических осей тоже две. Они образуют угол друг с другом, который обычно зависит от длины волны. Анало-



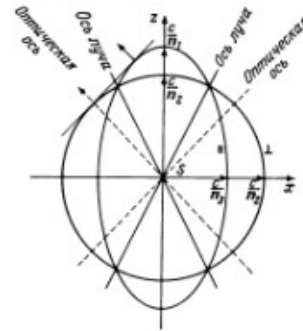
Ф и г. 15.4. Поверхность равной фазы для волны, распространяющейся из точечного источника S в двуосном кристалле.

гичное утверждение справедливо и для двух лучевых осей. Отметим, что одноосное двойное лучепреломление есть специальный случай двуосного, когда две оптические оси совпадают.

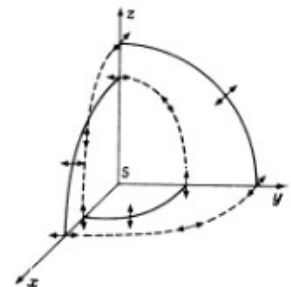
Наличие двух оптических осей в двуосном кристалле соответствует трем различным значениям показателя преломления. Эти три главных значения показателя преломления можно записать как $n_1 < n_2 < n_3$. Поверхность постоянной фазы приобретает сложную конфигурацию, показанную на фиг. 15.4. Каждая эквивалентная компонента образует круг в результате пересечения с одной из плоскостей и эллипсы — с двумя другими, перпендикулярными первой. В случае, который изображен на фиг. 15.4, оптические оси и оси лучей лежат в плоскости x, z (эта плоскость показана для ясности на фиг. 15.5). В плоскости x, z квазиобыкновенный луч распространяется с одинаковой скоростью c/n_2 вдоль любого радиуса-вектора, тогда как для необыкновенного луча существует зависимость скорости от направления распространения, которая графически представляется овалом. Отметим, что, как видно из фиг. 15.4, в двух других плоскостях, x, y и y, z , овал находится либо целиком внутри, либо целиком вне круга. Скорость распространения необыкновенного луча соответственно либо меньше, либо больше, чем у обыкновенного во всех направлениях, лежащих в рассматриваемой плоскости.

Характер поляризации волн иллюстрируется на фиг. 15.6, где показано, что квазиобыкновенный луч поляризован, причем вектор D перпендикулярен плоскости, в которой распространяется этот луч, тогда как необыкновенный луч поляризован в его собственной плоскости распространения.

В анизотропном кристалле электрическая индукция $D = \epsilon \mathcal{E}$ в общем случае не параллельна электрическому полю \mathcal{E} , поскольку диэлектрическая проницаемость ϵ представляет собой тензор. Электрическая поляризация, индуцированная вектором \mathcal{E}



Ф и г. 15.5. Сечение поверхности равной фазы плоскостью x, z , содержащей ось луча и оптическую ось, для волны, распространяющейся из центра S в двуосном кристалле.



Ф и г. 15.6. Поверхности равной фазы в двуосном кристалле. Стрелки указывают направление поляризации (электрического вектора) волн.

поля излучения, возникает в результате смещения зарядов в направлении D . Мы можем выбрать декартову систему координат таким образом, что вдоль трех осей смещения будут параллельны электрическому полю:

$$D_x = \epsilon_x \mathcal{E}_x, \quad D_y = \epsilon_y \mathcal{E}_y, \quad D_z = \epsilon_z \mathcal{E}_z. \quad (15.2)$$

Вспоминая соотношение между диэлектрической проницаемостью и показателем преломления, $\epsilon = n^2$, имеем

$$\epsilon_x = (n_1)^2, \quad \epsilon_y = (n_2)^2, \quad \epsilon_z = (n_3)^2. \quad (15.3)$$

Можно построить эллипсоид диэлектрической проницаемости (фиг. 15.7), определяемый формулой

$$\frac{x^2}{(n_1)^2} + \frac{y^2}{(n_2)^2} + \frac{z^2}{(n_3)^2} = 1. \quad (15.4)$$

Пересечение эллипсоида с плоскостью, проходящей через начало координат и перпендикулярной общему направлению распространения двух независимых плоских волн, представляет собой эллипс. Длины малой и большой полуосей этого эллипса равны показателям преломления n_a и n_b для этих двух волн. Две плоские

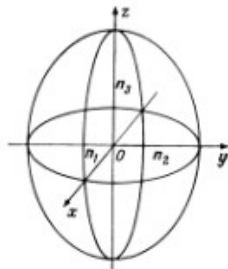
волны распространяются со скоростями соответственно c/n_a и c/n_b в направлении, перпендикулярном плоскости эллипса. Одна волна поляризована параллельно малой, другая — параллельно большой полуоси эллипса.

В общем случае направление распространения можно определить при помощи единичного вектора s с компонентами (в декартовой системе) s_x, s_y и s_z . Значения n_a и n_b можно получить, решая уравнение Френеля

$$\frac{(s_x)^2}{n^2 - (n_1)^2} + \frac{(s_y)^2}{n^2 - (n_2)^2} + \frac{(s_z)^2}{n^2 - (n_3)^2} = \frac{1}{n^2}. \quad (15.5)$$

Это уравнение (квадратичное относительно n) имеет два решения: $\pm n_a$ и $\pm n_b$. Знак \pm учитывает распространение волн в прямом и обратном направлениях.

До сих пор мы ограничивались специальным случаем, в котором два необыкновенных луча распространяются в одном и том же направлении s и отличаются только их взаимно перпендикулярными поляризациями и различными скоростями распространения вдоль s . В общем случае один необыкновен-



Ф и г. 15.7. Эллипсоид диэлектрической проницаемости двуосного кристалла.

ный луч распространяется в направлении s , другой — в направлении s' . Решение уравнения Френеля (15.5) для s и s' дает две пары показателей преломления n_a и n_b или n'_a и n'_b и соответственно две пары направлений поляризации. Каждому лучу можно приписать только одно значение n и одну поляризацию из каждой пары. Этот выбор определяется из рассмотрения граничных условий на той поверхности, где свет входит в кристалл.

Отметим, что в одноосном кристалле эллипсоид диэлектрической проницаемости обладает вращательной симметрией:

$$\frac{x^2}{(n_O)^2} + \frac{y^2}{(n_O)^2} + \frac{z^2}{(n_E)^2} = 1, \quad (15.6)$$

где z — оптическая ось.

До сих пор мы считали, что кристалл совершенно прозрачен для различных лучей. Фактически в полупроводнике имеется копочное поглощение, которое может быть различным для света, распространяющегося вдоль разных осей. Это свойство называется плеохроизмом. В дихроических кристаллах одна из двух компонент с взаимно перпендикулярными поляризациями поглощается

сильнее чем другая; следовательно, свет, прошедший через кристалл, линейно поляризован. Трихромические кристаллы обладают различными коэффициентами поглощения для света, поляризованного вдоль каждой из трех осей.

§ 2. ИНДУЦИРОВАННАЯ ОПТИЧЕСКАЯ АНИЗОТРОПИЯ

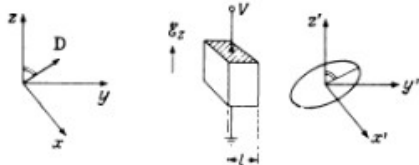
Циркулярно поляризованное излучение представляет собой волну, вектор D которой вращается относительно направления распространения. Возможно вращение как по часовой, так и против часовой стрелки. Кристалл называется оптически активным, если волны, циркулярно поляризованные по часовой и против часовой стрелки, распространяются с разной скоростью. Некоторые двоякопреломляющие вещества обладают этим свойством, будучи в естественном состоянии, другие становятся оптически активными под действием магнитного поля.

Отметим, что линейно поляризованное излучение можно разложить на две циркулярно поляризованные (по часовой и против часовой стрелки) компоненты, распространяющиеся с одинаковой скоростью. Следовательно, когда линейно поляризованный свет проходит через оптически активный материал, плоскость поляризации постепенно поворачивается.

Многие полупроводники, оптические свойства которых в обычных условиях изотропны, становятся анизотропными под действием электрического и магнитного полей, а также механического напряжения.

1. Электрооптический эффект Керра

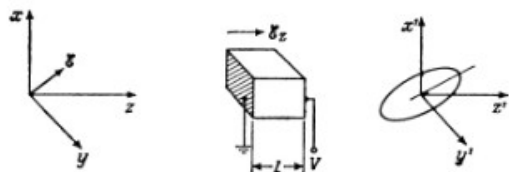
В электрооптическом эффекте Керра электрическое поле, приложенное к кристаллу в направлении, перпендикулярном направ-



Фиг. 15.8. Эффект Керра для линейно поляризованной волны. После прохождения через электрооптический кристалл свет становится эллиптически поляризованным.

лению распространения света, индуцирует двойное лучепреломление; при определенной кристаллографической ориентации кристалл становится одноосным с оптической осью, параллельной электрическому полю (фиг. 15.8). В общем случае полупро-

Отметим, что $E_z l$ есть напряжение, приложенное к кристаллу. Следовательно, для света, распространяющегося в направлении



Фиг. 15.9. Эффект Погкельса. Превращение линейно поляризованного света в эллиптически поляризованный.

приложенного поля, наблюдается линейный электрооптический эффект, величина которого для данного материала зависит только от приложенного напряжения.

Значения электрооптических коэффициентов γ при комнатной температуре для некоторых полупроводников приведены в табл. 15.1.

Таблица 15.1

Электрооптические свойства полупроводников

Вещество	Электрооптический коэффициент γ , 10^{-10} см/В	Показатель преломления при нулевом поле n_0	$n_0^3 \gamma$, 10^{-10} см/В	Литература
GaAs	1,6	3,34	59	[3]
ZnTe	1,4	2,79	34	[3]
ZnS	2,0	2,37	27	[4]
CdTe	6,8	2,6	120	[5]
CdS	5,5	2,3	66	[3]
Se	2,5	2,8	55	[6]

водник становится двуосным [2]. Если электрический вектор линейно поляризованного света образует угол 45° с направлением приложенного электрического поля, то компоненты этого вектора, направленные параллельно и перпендикулярно полю, распространяются с различными скоростями, так что свет выходит из кристалла поляризованным эллиптически. Оптическая разность хода для двух компонент определяется соотношением

$$\Delta = c_K l (E_z)^2 \lambda, \quad (15.7)$$

где c_K — постоянная Керра и l — длина (в направлении распространения света) той области кристалла, где есть электрическое поле.

2. Эффект Погкельса или линейный электрооптический эффект

В некоторых полупроводниках, которые в обычных условиях изотропны, электрическое поле вызывает двойное двойное лучепреломление. Для того чтобы понять влияние электрического поля на показатель преломления, рассмотрим эллипсоид диэлектрической проницаемости (фиг. 15.7). Когда электрическое поле равно нулю, показатель преломления изотропен и этот эллипсоид представляет собой сферу. Если электрическое поле приложено вдоль некоторого направления, которое мы обозначим через z , показатель преломления n_z остается постоянным. Однако в плоскости, перпендикулярной z , показатель преломления изменится на величину Δn , пропорциональную электрическому полю E_z . Сфера диэлектрической проницаемости преобразуется таким образом, что показатель преломления изменится на величину $\pm \Delta n$ во взаимно перпендикулярных направлениях. Тогда мы можем так выбрать два направления x и y , что

$$\begin{aligned} n_x &= n + \Delta n, \\ n_y &= n - \Delta n. \end{aligned} \quad (15.8)$$

При определенной ориентации электрического поля по отношению к кристаллографическим осям эффект оказывается максимальным. Выбор этой ориентации зависит от симметрии кристалла.

Величина линейного электрооптического эффекта определяется соотношением

$$\Delta n = \frac{n^3}{2} r E_z, \quad (15.9)$$

где r — электрооптический коэффициент.

В случае, представленном на фиг. 15.9, две компоненты, на которые можно разложить линейно поляризованный свет, распространяются с различными скоростями, что приводит к разности фаз

$$\Delta \phi = \frac{\pi}{\lambda} n^3 r E_z l. \quad (15.10)$$

2.12. ЦИКЛОТРОННЫЙ РЕЗОНАНС

Существует несколько экспериментальных методов определения эффективной массы носителей заряда. Одним из них является метод циклотронного резонанса. В этом случае полупроводник помещают в постоянное магнитное поле с индукцией $B = \text{const}$. На

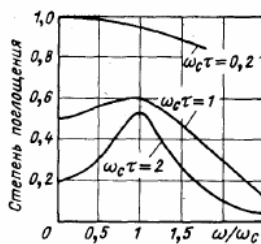


Рис. 2.18. Поглощение при циклотронном резонансе в зависимости от напряженности постоянного магнитного поля (в единицах ω/ω_c).

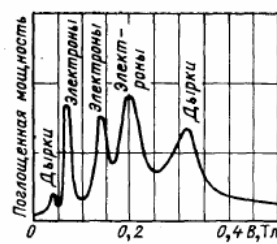


Рис. 2.19. Поглощение при циклотронном резонансе в германии (B — индукция)

электрон, движущийся со скоростью v , будет действовать сила Лоренца

$$F = -e[vB]. \quad (2.148)$$

Под действием этой силы электрон будет двигаться по окружности в плоскости, перпендикулярной магнитному полю. Круговая частота вращения ω_c , которая носит название ц и к л о т р о н н о й, определяется из равенства центробежной силы и силы Лоренца

$$m^* v^2 / r = e v B \quad (2.149)$$

и выражается формулой

$$\omega_c = \frac{v}{r} = \frac{eB}{m^*}. \quad (2.150)$$

Отсюда следует, что циклотронная частота ω_c не зависит от скорости и радиуса окружности. Если теперь в полупроводнике создать слабое высокочастотное электрическое поле, колеблющееся в плоскости, перпендикулярной \mathbf{B} , то при приближении частоты поля ω к ω_c электрон будет сильно поглощать энергию электромагнитного поля. Это явление резонансного поглощения принято называть циклотронным резонансом. Из измерений резонансной частоты поглощения радиоизлучения определяется ω_c , а затем по (2.150) вычисляется эффективная масса электрона (или дырки). Если эта величина анизотропна, то, изменяя ориентацию магнитного и электрического полей относительно кристаллографических осей кристалла, можно изучить форму эллипсоида равной энергии.

Осуществить эксперимент для определения резонансной частоты поглощения ω_c возможно в том случае, если время свободного пробега электрона τ достаточно велико:

$$\tau \gg \frac{1}{\omega_c}, \quad (2.151)$$

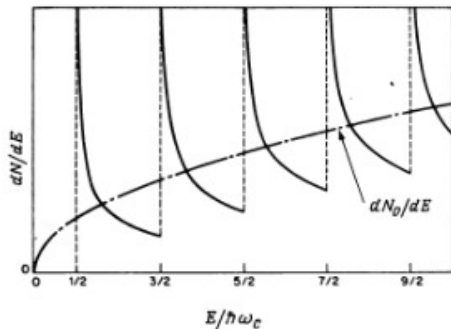
т. е. за время τ электрон должен успеть совершить хотя бы один оборот.

Чем больше электрон делает оборотов, тем резче, острее проявляется резонанс. На рис. 2.18 изображены кривые поглощения электромагнитного поля в зависимости от ω/ω_c для трех значений: $\omega_c\tau = 0,2$; 1 и 2.

Время свободного пробега τ носителя заряда тем больше, чем меньше в кристалле структурных несовершенств. Поэтому эксперименты по циклотронному резонансу необходимо проводить при очень низких температурах (обычно при 4,2 К — температуре жидкого гелия) и на особо чистых кристаллах. Обычно в опытах задается частота радиоизлучения, а резонанс достигается изменением индукции магнитного поля. На рис. 2.19 приведен спектр поглощения при циклотронном резонансе в германии.

где g — гиромангнитное отношение (значения m^* и g для различных соединений III—V и II—VI приведены в табл. 2.4 и 2.5).

Отметим, что магнитное поле сдвигает дно зоны проводимости вверх на величину $(1/2)q\hbar H/m^*c$. Сдвиг вершины валентной зоны



Фиг. 2.4. Плотность состояний в магнитном поле и без поля (штрих-пунктирная линия) [24].

в сторону меньших энергий составляет $(1/2)q\hbar H/m^*c$. Изменение ширины запрещенной зоны дается формулой

$$\Delta E_g = \frac{1}{2} \left(\frac{1}{m_c^*} + \frac{1}{m_v^*} \right) \frac{q\hbar H}{c}.$$

Следует также отметить, что зона проводимости и валентная зона состоят из нескольких долин, каждой из которых соответствует своя система дискретных энергетических уровней. Более того, поскольку величина эффективной массы зависит от направления, эти уровни могут быть уширены (примеси и тепловые колебания вызывают дополнительное уширение). Зависимость размеров циклотронной орбиты от направления была использована для определения эффективных масс носителей, соответствующих различным кристаллографическим направлениям, и выяснения таким образом формы долин в пространстве волновых векторов. Детальное рассмотрение циклотронного резонанса и обзор работ были выполнены Ляксом и Мавроидесом [27].

1. Расщепление Ландау

Если полупроводник помещен в магнитное поле H_z , направленное вдоль оси z , то это не оказывает влияния на движение электрона вдоль этой оси, однако в плоскости, перпендикулярной полю, возникает периодическое круговое движение с угловой частотой:

$$\omega_c = \frac{qH_z}{m^*c};$$

частота ω_c называется циклотронной частотой. Полученные орбиты квантованы, и разрешенные значения энергии (соответствующие движению в направлении, перпендикулярном полю) даются выражением

$$E_{xy} = \frac{q\hbar H_z}{m^*c} \left(n + \frac{1}{2} \right),$$

где n — целое число ($n \geq 0$). В продольном направлении энергия электрона в параболической зоне не зависит от магнитного поля и по-прежнему ее можно записать в виде

$$E_z = \frac{\hbar^2 k^2}{2m^*}.$$

Однако подсчет числа состояний, относящихся к различным энергиям [23, 24], показывает, что теперь вместо распределения, описываемого корневой зависимостью, состояния оказываются сконцентрированными в дискретные подзоны, как показано на фиг. 2.4. Плотность состояний теперь выражается формулой [23]

$$\frac{dN}{dE} = \left(\frac{2m^*}{\hbar^2} \right)^{3/2} \frac{1}{(2\pi l)^2} \sum_{n, \pm} \left[E - \left(n + \frac{1}{2} \pm \frac{v}{2} \right) \hbar\omega_c \right]^{-1/2},$$

где l — радиус наименьшей циклотронной орбиты, определяемой соотношением

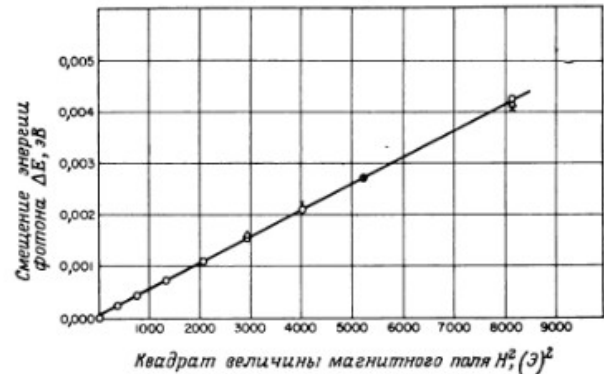
$$\frac{1}{l^2} = \frac{m^*}{\hbar} \omega_c,$$

а

$$v = \frac{m^*g}{2m},$$

2. Эффект Зеемана ¹⁾

Магнитное поле действует на электрон, движущийся по орбите, и вызывает расщепление уровня, который он занимал в отсутствие



Фиг. 2.5. Зависимость смещения энергии излучаемого фотона в GaAs от магнитного поля при 4,2 К [29].

Это смещение было приписано влиянию магнитного поля на энергетический уровень водородоподобного донора.

поля. Это расщепление, соответствующее ларморовской прецессии орбиты, при малых полях составляет $\pm (q\hbar H/2m^*c)$, т. е. половину расщепления Ландау. В случае сильных магнитных полей и больших орбит расщепление растет пропорционально H^2 . Эта нелинейность возникает, когда H превышает внутреннее поле, обусловленное совместным действием спина и орбитального движения. Экситоны и примеси часто характеризуются большими размерами орбит. Поэтому у примесей и экситонов наблюдается квадратичный эффект Зеемана (фиг. 2.5). У состояний разрешенных зон, напротив, наблюдается большее по величине, но линейное смещение.

Задача 1. В эффекте Келдыша — Франца вероятность $P(E)$ найти в пределах запрещенной зоны электрон с энергией E (отсчитанной от края зоны проводимости E_c) пропорциональна

$$\exp \left(- \frac{E - E_c}{\Delta E} \right)^{3/2},$$

где ΔE выражается формулой (2.4).

Нарисуйте край зоны в присутствии электрического поля и выразите упомянутую вероятность через поле \mathcal{E} и координату x . Разрешите кажущиеся противоречия между $P(x, \mathcal{E})$ и $P(E)$, где последняя величина показывает, что вероятность найти электрон ниже края зоны проводимости увеличивается с ростом поля.

3. Эффект Фарадея

Некоторые материалы становятся оптически активными в магнитном поле. Когда линейно поляризованное излучение проходит через материал в направлении, параллельном магнитному полю, плоскость поляризации поворачивается [7]. Угол поворота θ пропорционален магнитному полю H и длине кристалла l :

$$\theta = C_V H l.$$

Здесь C_V — так называемый коэффициент Верде (поворот на единице длины в единичном магнитном поле). Коэффициент Верде меняется с длиной волны. Направление вращения плоскости поляризации зависит от полярности магнитного поля. Принято считать положительным вращение, соответствующее повороту правого винта, перемещаемого в направлении магнитного поля. Отметим, что направление вращения одинаково для волны, распространяющейся в любом направлении вдоль поля.

Свободные носители вызывают эффект Фарадея, который легко наблюдать в сильно легированных полупроводниках. Величина фарадеевского вращения θ , вызванного свободными носителями, определяется выражением [8]

$$\theta = \frac{q^2 N \lambda^2 H l}{2 \pi c^4 n (m^*)^2}, \quad (15.11)$$

где N — концентрация свободных носителей, λ — длина волны излучения, n — показатель преломления в отсутствие магнитного поля, m^* — эффективная масса свободных носителей. Хотя в выражении (15.11) подразумевается, что эффективная масса изотропна, можно внести поправки, учитывающие эллипсоидальность долин [9] и наличие носителей двух типов.

Квадратичная зависимость θ от λ указывает, что эффект Фарадея на свободных носителях следует измерять в области длинных волн. Однако частота света должна быть велика по сравнению с циклотронной частотой $\omega_c = qH/m^*$, и, более того, соответствующий период должен быть мал по сравнению со временем релаксации носителей.

Если концентрация носителей N известна из других измерений, то эффект Фарадея можно использовать для определения эффективной массы.

§ 1. Неравновесные носители заряда

В полупроводниках, в отличие от металлов, под влиянием внешних воздействий (освещения, электрического тока в неоднородных структурах и других) концентрации электронов и дырок могут изменяться на много порядков. Это приводит к ряду специфических явлений, которые лежат в основе действия многих полупроводниковых приборов.

Чтобы пояснить, почему изменяются концентрации электронов и дырок, рассмотрим группу состояний электронов, которые изображаются точками в p -пространстве, заключенными внутри малого объема δp_1 (рис. 7.1). Часть из этих состояний фактически заполнена электронами, которые под влиянием теплового движения (сюда включается и тепловое излучение, находящееся в равновесии с кристаллом) могут переходить в другие состояния, например, принадлежащие объему δp_2 . При этом энергия электронов может изменяться. Обозначим через v_{12} темп таких переходов, т. е. число переходов в единицу времени, рассчитанное на единицу объема кристалла, а темп обратных переходов из состояния 2 в состояние 1 — через v_{21} . Тогда при термодинамическом равновесии устанавливается такое распределение электронов по квантовым состояниям, что

$$v_{12} = v_{21}. \quad (1.1)$$

Суммарные концентрации электронов и дырок равны при этом их равновесным значениям n_0 и p_0 . Соотношение (1.1) выражает принцип *детального равновесия*, который играет большую роль при статистических расчетах и более подробно будет рассмотрен в гл. XIII.

Существенно подчеркнуть, что взаимное уравновешивание прямых и обратных переходов справедливо не только для всей совокупности состояний, но и для отдельных их физически малых групп. Так, например, рассматривая тепловые возбуждения электронов

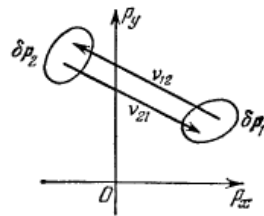


Рис. 7.1. Прямые и обратные электронные переходы.

4. Эффект Фойгта

Магнитное поле вызывает двойное лучепреломление, когда излучение распространяется в направлении, перпендикулярном магнитному полю. В этом случае мы имеем дело с магнитооптическим аналогом эффекта Керра. Если излучение линейно поляризовано, то компоненты волнового вектора \vec{k} , направленные перпендикулярно и параллельно магнитному полю, распространяются с различной скоростью, в результате излучение становится эллиптически поляризованным.

Эффект Фойгта, связанный со свободными носителями, находящимися в эллипсоидальных долинах, вызывает следующий сдвиг фаз [10]:

$$\Delta\phi = \frac{q^2 N \lambda^2 H^2 l}{4 \pi^2 c^4 n (m^*)^3},$$

где N — концентрация носителей, l — толщина образца, n — показатель преломления при нулевом поле и m^* — эффективная масса Фойгта, представляющая собой результат сложного усреднения эффективных масс, характеризующих эллипсоидальную долину. Величина m^* сильно зависит от кристаллографического направления.

из валентной зоны в зону проводимости, можно непосредственно заключить, что темп этих переходов (тепловой генерации электронно-дырочных пар) равен темпу обратных выпадений электронов из зоны проводимости в валентную зону (или рекомбинации электронно-дырочных пар), что вытекает просто из условия постоянства концентраций электронов и дырок в зонах. Однако согласно принципу *детального равновесия* можно утверждать гораздо больше, а именно что это равенство справедливо и для отдельных групп энергетических уровней в зонах (рис. 7.2).

При наличии внешних воздействий на полупроводник к тепловым переходам v_{12} добавляются переходы v'_{12} нетепловой природы, и при этом частота обратных переходов v_{21} тоже изменяется. Состояние полупроводника в таких условиях мы будем называть *термодинамически неравновесным*. При этом принцип *детального равновесия*, вообще говоря, уже не выполняется.

При нарушении термодинамического равновесия концентрации электронов и дырок в зонах n и p изменяются по сравнению с их равновесными значениями n_0 и p_0 , т. е. в зонах появляются *неравновесные* носители заряда с концентрациями $\delta n = n - n_0$ и $\delta p = p - p_0$. При этом изменяются и концентрации связанных носителей n_t и p_t .

Сказанное справедливо, конечно, и для металлов. Однако концентрация электронов в металлах гораздо больше, чем в полупроводниках, и ее относительное изменение обычно ничтожно мало.

Отметим, что уже в гл. VI мы имели дело, по существу, с неравновесными состояниями, так как наличие тока нарушает термодинамическое равновесие (ср. § VI. 3). Однако там можно было считать, что по отношению к концентрации носителей заряда равновесие сохраняется. Это значит, что для данного полупроводника при заданной температуре установившаяся концентрация электронов определяется только значением электростатического потенциала (ср., например, § VI.12). Теперь мы рассмотрим явления, в которых нарушается равновесие и по концентрациям. Именно такие состояния мы будем называть в дальнейшем *неравновесными*.

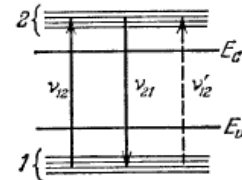


Рис. 7.2. К иллюстрации принципа *детального равновесия*. Тепловая генерация электронно-дырочных пар ($1 \rightarrow 2$) и обратный процесс рекомбинации электронов и дырок ($2 \rightarrow 1$).

§ 2. Время жизни неравновесных носителей заряда

В настоящей главе мы будем рассматривать поведение неравновесных носителей заряда только в объеме полупроводника, отвлекаясь от возможного влияния его поверхности. Это можно сделать, если отношение поверхности к объему образца достаточно мало (строгая говоря, для бесконечно протяженного полупроводника). Однако основные понятия и соотношения, вводимые в настоящей главе, сохраняются и в тех случаях, когда процессами генерации и рекомбинации на самой поверхности пренебрегать нельзя. Влияние поверхности на неравновесные состояния будет рассмотрено дополнительно в гл. X, где будут выяснены и условия, при которых этим влиянием можно пренебречь.

Положим, что под влиянием внешнего воздействия в каждой единице объема полупроводника в единицу времени возникает g_n электронов проводимости и, соответственно, g_p дырок в валентной зоне. Темпы генерации g_n и g_p будем считать сначала одинаковыми по всему объему, хотя и не обязательно равными друг другу. Пусть, далее, R_n есть темп обратного процесса исчезновения свободных электронов вследствие процессов рекомбинации с дырками (свободными и связанными на локальных уровнях энергии) и, соответственно, R_p — темп рекомбинации свободных дырок. Если в полупроводнике нет электрического тока, то изменение во времени неравновесных концентраций электронов и дырок в зонах определяется уравнениями

$$\frac{dn}{dt} = g_n - R_n, \quad \frac{dp}{dt} = g_p - R_p. \quad (2.1)$$

Подчеркнем, что g_n и g_p обозначают генерацию, обусловленную только *внешними* воздействиями, и не включают переходы, вызванные тепловым движением. Последние мы учитываем в величинах R_n и R_p , которые поэтому представляют результирующие темпы рекомбинации, т. е. разности между темпом захвата носителей из соответствующей зоны и темпом обратной тепловой генерации носителей в зону:

$$R_n = r_n - g_{nT}, \quad R_p = r_p - g_{pT}. \quad (2.2)$$

Для количественного описания кинетики неравновесных электронных процессов широко применяют понятия *среднего времени жизни неравновесных электронов* в зоне проводимости τ_n и, соответственно, *дырок* в валентной зоне τ_p , которые определяются формулами

$$R_n = \frac{n - n_0}{\tau_n}, \quad R_p = \frac{p - p_0}{\tau_p}. \quad (2.3)$$

Или иначе: $1/\tau_n$ есть вероятность исчезновения одного избыточного электрона из зоны проводимости в единицу времени вследствие рекомбинации (со свободными и связанными дырками). Аналогично, $1/\tau_p$ есть вероятность рекомбинации одной избыточной дырки, тоже за единицу времени.

Так как r_n и r_p в формулах (2.2) по физическому смыслу не могут иметь слагаемых, не зависящих от n и, соответственно, p , то из

сравнения формул (2.2) и (2.3) следует, что

$$g_{nT} = \frac{n_0}{\tau_n}, \quad g_{pT} = \frac{p_0}{\tau_p}. \quad (2.4)$$

Следовательно, времена τ_n и τ_p определяют не только темп суммарной рекомбинации, но и темп тепловой генерации электронов и дырок.

Пользуясь понятиями времен жизни, уравнения кинетики (2.2) для однородного образца без тока можно записать в виде

$$\frac{dn}{dt} = g_n - \frac{n}{\tau_n}, \quad \frac{dp}{dt} = g_p - \frac{p}{\tau_p}. \quad (2.5)$$

Стационарные концентрации неравновесных носителей заряда, устанавливающиеся после длительного воздействия внешней генерации, равны

$$(\delta n)_s = g_n \tau_n, \quad (\delta p)_s = g_p \tau_p. \quad (2.6)$$

Величины τ_n и τ_p зависят от физических особенностей элементарных актов рекомбинации электронов и дырок. При этом τ_n и τ_p , вообще говоря, могут сами зависеть от неравновесных концентраций δn и δp , а также от температуры. Поэтому τ_n и τ_p не являются характеристиками данного полупроводника, но зависят еще от условий опыта. Эти вопросы будут подробнее рассмотрены в гл. IX (см. также § XVII.9). Сейчас же мы будем считать τ_n и τ_p заданными феноменологическими величинами, определяющими кинетику электронных процессов.

В простейшем случае, когда τ_n и τ_p не зависят от n и p , интегрирование уравнений (2.5) дает

$$\delta n = g_n \tau_n - C \exp(-t/\tau_n)$$

и такое же выражение для концентрации неравновесных дырок δp . Здесь C — постоянная интегрирования, определяемая начальными условиями. Так, если вначале полупроводник находился в термодинамическом равновесии и затем в момент времени $t = 0$ было включено внешнее воздействие, мы имеем $t = 0$, $\delta n = 0$. Это дает

$$C = g_n \tau_n = (\delta n)_s, \quad \delta n = (\delta n)_s [1 - \exp(-t/\tau_n)]. \quad (2.7a)$$

Если в некоторый момент времени $t = t_1$ генерация выключается, то для $t \geq t_1$, $g_n = 0$, а начальное условие есть $t = t_1$, $\delta n = (\delta n)_1$. Тогда

$$C = -(\delta n)_1 \exp\left(\frac{t_1}{\tau_n}\right), \quad \delta n = (\delta n)_1 \exp\left(-\frac{t - t_1}{\tau_n}\right). \quad (2.7b)$$

При постоянном τ_n установление избыточной концентрации электронов и ее исчезновение описываются экспоненциальным законом, а среднее время жизни τ_n есть время, в течение которого концентрация неравновесных электронов изменяется в e раз. То же справедливо и для неравновесных дырок.

В общем случае τ_n и τ_p , изменяющихся вместе с n и p , соотношения (2.3) определяют *мгновенные* времена жизни. Величины $1/\tau_n$ и $1/\tau_p$ в этом случае дают вероятности рекомбинации одной частицы в единицу времени при данных значениях n и p , а τ_n и τ_p определяют времена, за которые неравновесные концентрации электронов или, соответственно, дырок изменились бы в e раз, если бы вероятность их рекомбинации сохранялась постоянной и равной данному мгновенному значению.

§ 5. Квазиуровни Ферми

В термодинамически неравновесном состоянии (например, при освещении полупроводника) уже не существует единого уровня Ферми для всей системы и поэтому выражения для концентраций электронов и дырок, полученные нами в гл. V, уже несправедливы. При этом, в частности, не выполняется и соотношение $np = n_i^2$.

Однако, следуя Шокли, можно обобщить соотношения статистики на неравновесные состояния, если вместо уровня Ферми формально ввести новые величины — *квазиуровни Ферми*. Положим, что вероятность заполнения электроном состояния с энергией E в зоне проводимости можно представить в виде, по форме совпадающем с распределением Ферми — Дирака (V.3.1):

$$f_n = \left(1 + \exp\left(\frac{E - F_n}{kT}\right)\right)^{-1}. \quad (5.1)$$

Тогда, по определению, величина F_n есть квазиуровень Ферми для электронов.

Аналогично, для вероятности нахождения вакансии (дырки) на уровне энергии E в валентной зоне положим

$$f_p = \left(1 + \exp\left(\frac{F_p - E}{kT}\right)\right)^{-1}, \quad (5.2)$$

где, по определению, F_p есть квазиуровень Ферми для дырок. Тогда очевидно, что для n и p мы получим те же соотношения, что и в гл. V, в которые, однако, вместо уровня Ферми F будут входить квазиуровни F_n и, соответственно, F_p . В частности, в невырожденных полупроводниках для неравновесных электронов будет по-прежнему справедливо распределение Больцмана

$$f_n = \exp\left(\frac{F_n - E}{kT}\right). \quad (5.1a)$$

Соответственно вместо формул (V.5.1) и (V.5.2) мы получим

$$n = n_0 + \delta n = N_c \exp\left(\frac{F_n - E_c}{kT}\right), \quad (5.3)$$

$$p = p_0 + \delta p = N_v \exp\left(\frac{E_v - F_p}{kT}\right). \quad (5.4)$$

Вместо формулы (V.5.4) мы будем иметь

$$np = n_i^2 \exp\left(\frac{F_n - F_p}{kT}\right). \quad (5.5)$$

Таким образом, появление в зонах неравновесных электронов и дырок можно описать как «расщепление» первоначального уровня Ферми F на два квазиуровня F_n и F_p , каждый из которых смещается по направлению к своей зоне (рис. 7.6). Совершенно аналогично, для нахождения концентраций связанных носителей на ловушках n_l и p_l мы могли бы воспользоваться функцией распределения (V.9.3) (или более общей формулой (V.11.8)), заменив в ней F на некоторый квазиуровень Ферми для ловушек F_l , который, вообще говоря, отличается от F_n и F_p .

Введение квазиуровней Ферми физически обозначает предположение, что времена релаксации импульса τ_p и энергии τ_E для электронов и дырок намного меньше времени их существования в зонах. Как уже говорилось в § 4, в этом случае можно считать, что в электронном и дырочном газах устанавливается равновесное фермиевское распределение, и притом с одной и той же температурой для всей системы. Однако равновесия по отношению к концентрациям электронного и дырочного газов при этом может и не быть. Именно это и учитывается введением различных квазиуровней Ферми для дырок и электронов. Отметим, что указанное предположение а priori не очевидно. Однако имеются эксперименты, позволяющие его проверить (см. § 9). Они показывают, что, по крайней мере для некоторых полупроводников, оно действительно оправдывается.

Отметим теперь два важных свойства квазиуровней Ферми. В § VI.3 мы видели, что плотность тока в невырожденных полупроводниках пропорциональна градиенту уровня Ферми. Повторяя те же рассуждения для рассматриваемого случая биполярной неравновесной проводимости и используя вместо соотношений (V.5.1) и (V.5.2) обобщенные формулы (5.3) и (5.4), мы получим для электронной и дырочной составляющих плотности тока j_n и j_p формулы, аналогичные (VI.3.2):

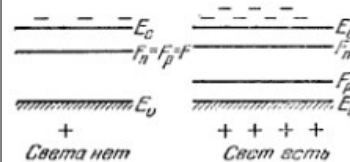


Рис. 7.6. К понятию квазиуровней Ферми.

Именно это и учитывается введением различных квазиуровней Ферми для дырок и электронов. Отметим, что указанное предположение а priori не очевидно. Однако имеются эксперименты, позволяющие его проверить (см. § 9). Они показывают, что, по крайней мере для некоторых полупроводников, оно действительно оправдывается.

Отметим теперь два важных свойства квазиуровней Ферми. В § VI.3 мы видели, что плотность тока в невырожденных полупроводниках пропорциональна градиенту уровня Ферми. Повторяя те же рассуждения для рассматриваемого случая биполярной неравновесной проводимости и используя вместо соотношений (V.5.1) и (V.5.2) обобщенные формулы (5.3) и (5.4), мы получим для электронной и дырочной составляющих плотности тока j_n и j_p формулы, аналогичные (VI.3.2):

$$j_n = \mu_n n(r) \nabla F_n, \quad (5.6)$$

$$j_p = \mu_p p(r) \nabla F_p. \quad (5.7)$$

Отсюда видно, что при наличии тока квазиуровни Ферми изменяются в пространстве, и тем быстрее, чем меньше локальные значения концентраций электронов n или, соответственно, дырок p .

Далее, легко видеть, что разность квазиуровней Ферми на концах полупроводника непосредственно связана с разностью электрических потенциалов, существующей между этими концами. Положим сначала, что в полупроводнике имеются только электроны и будем считать для простоты, что все величины зависят от одной координаты x . Тогда, вводя в формулу (5.6) электропроводность $\sigma_n(x) = e\mu_n n(x)$, имеем

$$j_{nx} = \frac{1}{e} \sigma_n(x) \frac{dF_n}{dx}.$$

С другой стороны, согласно общей форме дифференциального закона Ома, можно написать

$$j_{nx} = \sigma_n(x) (\mathcal{E}_x + \mathcal{E}_{nx}^*),$$

где \mathcal{E}_x — напряженность кулоновского электрического поля, а \mathcal{E}_{nx}^* — напряженность поля сторонних сил, действующих на электроны. (В рассматриваемом случае это есть сила давления электронного газа, обусловленная градиентом концентрации электронов). Сравнивая эти соотношения, получаем

$$\frac{dF_n}{dx} = e (\mathcal{E}_x + \mathcal{E}_{nx}^*).$$

Поэтому разность значений квазиуровней Ферми для двух каких-либо сечений полупроводника A и B (рис. 7.7) равна

$$F_{nB} - F_{nA} = e \left(\int_A^B \mathcal{E}_x dx + \int_A^B \mathcal{E}_{nx}^* dx \right).$$

Первый интеграл в правой части есть уменьшение потенциала при переходе от A к B , обусловленное током (или i , где i — сила тока, а r — сопротивление участка). Второй интеграл есть электродвижущая сила V_0 . Она вызывает увеличение потенциала при переходе от A к B на V_0 . Поэтому все выражение в круглых скобках есть

$$ir - V_0 = \text{разности потенциалов на концах участка } (\varphi_A - \varphi_B)$$

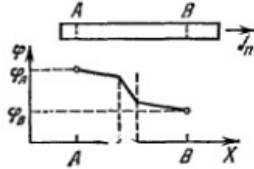


Рис. 7.7. К установлению связи между квазиуровнями Ферми и внешним напряжением.

§ 1. Различные типы процессов рекомбинации

Рассматривая неравновесные электронные состояния в гл. VII, мы ввели понятия о средних временах жизни избыточных электронов и дырок в зонах τ_n и τ_p . Однако мы не интересовались тем, от чего зависят эти времена. Между тем опыт показывает, что τ_n и τ_p могут изменяться в чрезвычайно широких пределах — от многих часов до 10^{-8} с и меньше. В настоящей главе мы рассмотрим подробно процессы рекомбинации и генерации неравновесных электронов и дырок и выясним физические причины, влияющие на времена их жизни.

Процессы рекомбинации удобно разделить на два класса: 1) прямую рекомбинацию зона — зона и 2) рекомбинацию с участием примесей и дефектов. В первом случае свободный электрон из зоны проводимости рекомбинирует со свободной дыркой в валентной зоне в одном элементарном акте. В процессах второго класса свободные электроны рекомбинируют со связанными дырками на примесях или дефектах, а свободные дырки — со связанными электронами. Мы остановимся сначала на прямой рекомбинации зона — зона.

При переходе электрона из зоны проводимости в валентную зону всегда должны выполняться законы сохранения энергии и квазиимпульса. Поэтому если E' и k' — энергия и квазиволновой вектор электрона в начальном состоянии в зоне проводимости, а E и k — их значения в конечном состоянии в валентной зоне, то должны выполняться соотношения

$$E'(k') = E(k) + \Delta E, \quad (1.1)$$

$$\hbar k' = \hbar k + \Delta p. \quad (1.2)$$

Здесь ΔE и Δp — энергия, освобождаемая в элементарном акте рекомбинации и, соответственно, изменение квазиимпульса электрона при переходе. Поскольку квазиимпульсы в начальном и конечном состояниях системы должны лежать в пределах первой зоны Бриллюэна (§ III.3), правая часть уравнения (1.2) может содержать еще слагаемое $\hbar b$, которое мы не выписываем. Прямая рекомбинация

и

$$F_{pB} - F_{pA} = -e (\varphi_B - \varphi_A).$$

Распределение потенциала показано на рис. 7.7 снизу, где для ясности чертежа предположено, что здесь имеется только в некотором слое.

Рассуждая аналогично, мы найдем, что если бы были только дырки, то разность квазиуровней для дырок была бы

$$F_{pB} - F_{pA} = -e (\varphi_B - \varphi_A)',$$

где разность потенциалов теперь была бы, вообще говоря, другой, так как распределения неравновесных электронов и дырок различны и сторонние силы \mathcal{E}_{nx}^* и \mathcal{E}_{px}^* неодинаковы.

Выберем теперь сечения A и B в таких частях полупроводника, где концентрации избыточных носителей $\delta p = \delta n = 0$. Тогда $F_{nB} = F_{pB} = F_B$, $F_{nA} = F_{pA} = F_A$ и, следовательно, $(\varphi_B - \varphi_A) = (\varphi_B - \varphi_A)'$. Поэтому

$$F_B - F_A = -e (\varphi_B - \varphi_A), \quad (5.8)$$

где $(\varphi_B - \varphi_A)$ есть приложенное внешнее напряжение. Таким образом, разность квазиуровней Ферми есть та величина, которую мы непосредственно измеряем вольтметром.

ция электрона и дырки возможна только в тех случаях, когда от электрона может быть отведена освобождающаяся энергия ΔE и обеспечено изменение квазиимпульса Δp , определяемые законами сохранения (1.1) и (1.2).

Процессы, обеспечивающие выполнение законов сохранения, могут быть различны. Это может быть испускание кванта электромагнитного излучения — фотона определенной частоты ω , тогда *)

$$\Delta E = \hbar \omega, \quad |\Delta p| = \frac{\hbar \omega}{c}. \quad (1.3)$$

В этом случае мы говорим об *излучательной рекомбинации*. Избыточные энергия и квазиимпульс могут быть переданы колебаниям решетки, или фононам. Известны и процессы, при которых энергия и квазиимпульс, освобождаемые при рекомбинации, передаются третьей свободной частице — электрону (в электронных полупроводниках) или дырке (в дырочных полупроводниках). Процессы такого типа получили название *ударной рекомбинации* (или *Оже-рекомбинации*). Очевидно, что вероятность таких процессов возрастает при увеличении концентрации носителей заряда, и поэтому ударная рекомбинация проявляется обычно в сильно легированных полупроводниках. Принципиально возможны и другие типы элементарных актов рекомбинации (см. § XVII.9). В случаях, когда фотоны не участвуют в обмене энергией и квазиимпульсом, говорят о *безызлучательной рекомбинации*.

Вероятность различных типов элементарных актов зависит от энергетической структуры кристалла, а также от концентрации электронов и дырок. При этом процессы разного типа могут сосуществовать.

В настоящей главе мы будем рассматривать только статистику рекомбинации и получим соотношения, справедливые при любом механизме рекомбинации. Конкретная природа рекомбинационных процессов будет определять лишь значения фигурирующих в дальнейшем вероятностей электронных переходов (эффективных сечений рекомбинации).

§ 2. Темп рекомбинации зона — зона

Рассмотрим группу состояний 1 (рис. 9.1) в зоне проводимости с энергией в интервале $(E', E' + dE')$ и группу состояний 2 в валентной зоне с энергией $(E, E + dE)$. Число переходов $1 \rightarrow 2$, рассчитанное на единицу времени и единицу объема, должно быть пропорционально количеству электронов в состояниях 1, т. е. $N_c(E') dE' f(E')$, и количеству незаполненных состояний в группе 2, которое есть $N_v(E) dE f_p(E)$. Поэтому число таких переходов в единице объема и в единицу времени можно записать в виде

$$dr = W(E', E) N_c(E') \times N_v(E) f(E') f_p(E) dE' dE. \quad (2.1)$$

Здесь $W(E', E)$ — вероятность перехода, отнесенная к единице времени. В принципе ее можно вычислить методами квантовой механики (см. § XVIII.4). Чтобы найти полный темп рекомбинации, нужно просуммировать выражение (2.1) по всем энергиям E' и E .

Темп рекомбинации выражается особенно просто при выполнении двух условий: 1) существуют квазиуровни Ферми для электронов и дырок F_n и F_p и 2) полупроводник не вырожден. В этом случае (§ VII.5)

$$f(E') = \exp\left(\frac{F_n - E'}{kT}\right), \quad f_p(E) = \exp\left(\frac{E - F_p}{kT}\right), \quad np = n_0 p_0 \exp\left(\frac{F_n - F_p}{kT}\right),$$

где n и p — полные концентрации электронов и, соответственно, дырок, а n_0 и p_0 — их равновесные значения. Пользуясь еще соотношением (V.5.4), мы получаем

$$f(E') f_p(E) = \exp\left(\frac{F_n - F_p}{kT}\right) \cdot \exp\left(\frac{E - E'}{kT}\right) = \frac{np}{n_0 p_0} \exp\left(\frac{E - E'}{kT}\right) = \frac{np}{N_c N_v} \exp\left(-\frac{E' - E_c}{kT}\right) \cdot \exp\left(-\frac{E_v - E}{kT}\right). \quad (2.2)$$

Результат подстановки этого выражения в формулу (2.1) и последующего интегрирования по E' и E будет, очевидно, иметь вид

$$r = \alpha np, \quad (2.3)$$

где

$$\alpha = \frac{1}{N_c N_v} \int_{E' = E_c}^{\infty} \int_{E = -\infty}^{E_v} W(E', E) N_c(E') N_v(E) \times \exp\left(-\frac{E' - E_c}{kT}\right) \cdot \exp\left(-\frac{E_v - E}{kT}\right) dE' dE. \quad (2.4)$$

Поэтому для времени жизни неравновесных пар, по определению, мы получаем

$$\tau = \frac{\delta n}{R} = \frac{1}{\alpha(n_0 + p_0 + \delta n)}. \quad (2.5)$$

Таким образом, время жизни пар зависит от концентрации равновесных носителей n_0 , p_0 и концентрации неравновесных пар δn . Поэтому время жизни пар не является характеристикой данного типа полупроводника, а отражает только свойства конкретного образца при определенном его состоянии. Напротив, коэффициент рекомбинации или, соответственно, сечение рекомбинации являются характеристиками данного типа полупроводника. Однако они, конечно, зависят от температуры (ср. формулу (2.4)).

Отметим, наконец, что зависимость τ от n_0 и p_0 имеет резкий максимум. Учитывая, что равновесные концентрации n_0 и p_0 связаны соотношением $n_0 p_0 = n_i^2$, и приравняв производную от выражения (2.8) по n_0 (или p_0) нулю, мы находим, что τ достигает максимума при $n_0 = p_0$, т. е. в собственном полупроводнике. При малом уровне возбуждения ($\delta n < n_0 + p_0$) этот максимум равен

$$\tau_{\max} = \tau_i = \frac{1}{2\alpha n_i}. \quad (2.9)$$

Формула (2.3) показывает, что темп прямой рекомбинации зона — зона пропорционален произведению концентраций n и p , т. е. подчиняется тому же закону, что и скорость бимолекулярных химических реакций. Коэффициент α , по определению, есть коэффициент рекомбинации. Из формулы (2.4) видно, что он зависит от вероятности элементарного акта рекомбинации $W(E', E)$ и распределения квантовых состояний в зонах $N_c(E')$ и $N_v(E)$, т. е. от структуры энергетических зон. Кроме того, он зависит от температуры.

В гл. VII мы определили суммарный темп рекомбинации как разность

$$R = r - g_T = \alpha np - g_T,$$

где g_T есть темп тепловой генерации электронов и дырок в условиях равновесия. Так как в равновесии (при $n = n_0$ и $p = p_0$) $R = 0$, то

$$g_T = \alpha n_0 p_0 = \alpha n_i^2. \quad (2.5)$$

Поэтому мы получаем окончательно

$$R = \alpha (np - n_0 p_0). \quad (2.6)$$

Именно эта величина должна входить во введенные ранее уравнения непрерывности.

Вероятность элементарного акта рекомбинации часто выражают при помощи эффективного сечения рекомбинации. При этом, по аналогии с кинетической теорией газов, каждый акт рекомбинации электрона с дыркой чисто формально уподобляют акту соударения электрона и частицы с некоторым сечением S . Тогда число «соударений» электронов с одной дыркой за единицу времени будет равно, очевидно, числу электронов, заключенных внутри цилиндра с основанием S и длиной образующей, равной средней тепловой скорости электронов v_T (рис. 9.2), т. е. $S v_T n$. Полное число искомых элементарных актов с учетом не одной, а всех имеющихся дырок будет в p раз больше, т. е. $S v_T np$. Сравнивая это с выражением (2.3), мы имеем

$$\alpha = S v_T, \quad (2.7)$$

где S , по определению, есть эффективное сечение рекомбинации. При этом в определении (2.7) под v_T обычно понимают скорость изолированного электрона (с массой m_0).

Обратимся теперь к времени жизни. Так как при переходах зона — зона каждое появление электрона в зоне проводимости сопровождается появлением одной дырки в валентной зоне, то $\delta n = \delta p$. Полагая, далее, $n = n_0 + \delta n$, $p = p_0 + \delta n$, мы имеем

$$R = \alpha (np - n_0 p_0) = \alpha (p_0 + n_0 + \delta n) \delta n.$$

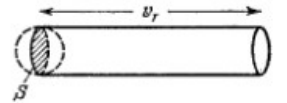


Рис. 9.2. К определению эффективного сечения рекомбинации.

§ 4. Рекомбинация через примеси и дефекты

Темп излучательной рекомбинации зона — зона велик в узкозонных полупроводниках. Она может быть преобладающей и в широкозонных полупроводниках с прямыми переходами (GaAs, CdS и др.). Однако очень часто в полупроводниках с широкой запрещенной зоной основную роль играют процессы рекомбинации с участием примесей и структурных дефектов. На это указывают многочисленные экспериментальные факты. Так, например, в § 3 мы видели, что если бы в кремнии излучательная рекомбинация была основным процессом, то время жизни пар должно было бы быть около 3 часов. Однако максимальные наблюдаемые времена жизни в кремнии при комнатных температурах не превышают нескольких миллисекунд. Опыт также показывает, что время жизни чрезвычайно чувствительно даже к ничтожному содержанию некоторых примесей. Если, например, в чистейшем кристалле германия n -типа с $\tau \sim 10^{-3}$ с ввести никель или золото (создающие глубокие локальные уровни энергии) с концентрацией $\sim 10^{15}$ см $^{-3}$ (т. е. всего 10^{-3} атомных %!), то время жизни электронов падает до $10^{-8} + 10^{-9}$ с, т. е. уменьшается на шесть порядков.

Процесс рекомбинации с участием локальных уровней (ловушек) является двухступенчатым и состоит из захвата электрона из зоны проводимости на пустой центр рекомбинации (ловушку) и захвата дырки из валентной зоны на заполненный центр рекомбинации. При этом сами центры рекомбинации не изменяются и играют роль своеобразного катализатора процесса рекомбинации.

Будем рассматривать сначала простые центры рекомбинации, которые могут захватывать или, соответственно, отдавать один-единственный электрон, а следовательно, различаться только в двух различных зарядных состояниях. Отвлекаясь от возбужденных

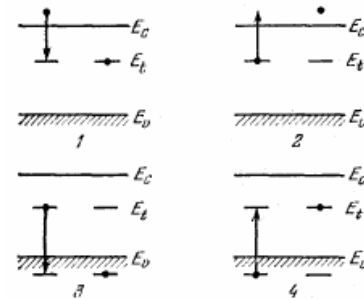


Рис. 9.6. Четыре типа элементарных процессов при рекомбинации электронов и дырок через локальные уровни.

состояний, такой центр можно охарактеризовать единственным локальным уровнем энергии для электронов E_i в запрещенной зоне (рис. 9.6). Если N_i есть полная концентрация ловушек (и равная ей концентрация локальных уровней), то в состоянии равновесия концентрация заполненных ловушек есть $N_i f_0$, а концентрация пустых равна $N_i(1 - f_0)$, где f_0 — функция Ферми. При нарушении равновесия концентрации заполненных и пустых ловушек изменяются и становятся $N_i f$ и $N_i(1 - f)$. Неравновесная вероятность заполнения ловушки f уже не выражается функцией Ферми и должна быть определена дополнительно (см. ниже).

В такой модели результирующая кинетика изменения концентрации электронов и дырок определяется четырьмя процессами: захватом электронов на ловушки (рис. 9.6, переходы 1), тепловой эмиссией электронов с ловушек в зону проводимости (переходы 2), захватом дырок на ловушки (переходы 3) и эмиссией дырок с ловушек в валентную зону (переходы 4). Найдем темпы каждого из этих переходов.

Рассмотрим группу электронов в зоне проводимости, энергия которых заключена в интервале $(E, E + dE)$. Число переходов $(E, E + dE) \rightarrow E_i$ в единице объема и в единицу времени пропорционально, во-первых, концентрации рассматриваемых электронов $N_c(E) dE \cdot f(E)$ и, во-вторых, концентрации пустых ловушек $N_i[1 - f(E_i)]$. Поэтому темп захвата электронов из зоны проводимости на ловушки есть

$$r_n = N_i [1 - f(E_i)] \int_{E_c}^{\infty} c_n(E) N_c(E) f(E) dE. \quad (4.1)$$

Здесь коэффициент $c_n(E)$ пропорционален отнесенной к единице времени вероятности элементарного акта перехода электрона с уровня E на уровень E_i . Выражение (4.1) гораздо удобнее переписать в виде

$$r_n = \alpha_n N_i [1 - f(E_i)] n, \quad (4.2)$$

где n — полная концентрация электронов в зоне, а α_n есть, по определению, коэффициент захвата электронов на ловушки. Сравнивая оба выражения для r_n и учитывая еще, что

$$n = \int_{E_c}^{\infty} N_c(E) f(E) dE,$$

мы видим, что

$$\alpha_n = \frac{\int_{E_c}^{\infty} c_n(E) N_c(E) f(E) dE}{\int_{E_c}^{\infty} N_c(E) f(E) dE}. \quad (4.3)$$

Или, иначе,

$$\alpha_n = \langle c_n \rangle, \quad (4.3a)$$

Рассуждения для дырок совершенно аналогичны. Отличие заключается в том, что темп захвата дырок из валентной зоны пропорционален концентрации ловушек, заполненных электронами (т. е. свободных от дырок), $N_i f$. Поэтому вместо выражения (4.2) мы будем иметь

$$r_p = \alpha_p N_i f p,$$

где α_p — коэффициент захвата для дырок. Соответственно темп тепловой реэмиссии дырок с ловушек будет

$$g_{pT} = \beta_p N_i (1 - f).$$

Приравнявая опять r_p и g_{pT} в случае равновесия, мы получим связь между β_p и α_p в виде

$$\beta_p = \alpha_p p_1,$$

где p_1 выражается формулой (V.9.13). Поэтому для суммарного темпа захвата дырок из валентной зоны на ловушки получаем

$$R_p = r_p - g_{pT} = \alpha_p N_i [p f - p_1 (1 - f)]. \quad (4.7a)$$

Обратимся теперь к процессам 2 (рис. 9.6) — тепловой эмиссии электронов с ловушек в зону. Очевидно, что темп таких переходов g_{nT} пропорционален концентрации ловушек, заполненных электронами, $N_i f(E_i)$. Поэтому можно положить

$$g_{nT} = \beta_n N_i f(E_i), \quad (4.5)$$

где β_n есть некоторый коэффициент, зависящий от природы ловушек и температуры.

Коэффициент реэмиссии β_n просто связан с коэффициентом захвата α_n . Эту связь можно найти с помощью принципа детального равновесия. Действительно, в состоянии термодинамического равновесия должно быть $r_n = g_{nT}$. Это дает

$$\alpha_n N_i (1 - f_0) n_0 = \beta_n N_i f_0, \quad (4.6)$$

где f_0 — равновесная вероятность заполнения ловушек электронами. Последнюю для наших целей удобно выразить не обычной функцией Ферми, а по формуле (V.9.4) записать в виде

$$f_0 = \frac{n_0}{n_0 + n_1}.$$

Здесь n_1 зависит от энергии ионизации ловушки $E_c - E_i$ и температуры и выражается формулой (V.9.5). После этого соотношение (4.6) дает

$$\beta_n = \alpha_n \frac{1 - f_0}{f_0} n_0 = \alpha_n n_1.$$

Следовательно, коэффициент реэмиссии вполне определяется через коэффициент захвата, глубину залегания энергетического уровня ловушек и, естественно, зависит еще от температуры. Подставляя найденное выражение для β_n в формулу (4.5), получаем

$$g_{nT} = \alpha_n n_1 \cdot N_i f.$$

Суммарный темп захвата электронов на ловушки равен

$$R_n = r_n - g_{nT} = \alpha_n N_i [n(1 - f) - n_1 f]. \quad (4.7)$$

Именно эти величины, R_n и R_p , мы должны подставлять в уравнение непрерывности (VII.3.3) для учета рекомбинации. Если изменения концентраций вследствие движения частиц малы, то эти уравнения принимают вид

$$\frac{dn}{dt} = g_n - R_n = g_n - \alpha_n N_i [n(1 - f) - n_1 f], \quad (4.8)$$

$$\frac{dp}{dt} = g_p - R_p = g_p - \alpha_p N_i [p f - p_1 (1 - f)]. \quad (4.9)$$

Здесь g_n и g_p — темпы внешней генерации электронов и, соответственно, дырок.

Так как при рекомбинации через ловушки каждый исчезающий свободный электрон оседает на ловушку (то же относится и к дыркам), то изменение концентрации заполненных ловушек $N_i f$ определяется уравнением

$$N_i \frac{df}{dt} = R_n - R_p + (g_p - g_n). \quad (4.10)$$

Кроме того, мы имеем еще уравнение квазинейтральности (ср. § VII.3)

$$p - p_0 = n - n_0 + N_i (f - f_0). \quad (4.11)$$

Уравнения (4.8) — (4.11) вполне определяют кинетику изменения концентрации неравновесных электронов и дырок вследствие рекомбинации.

Отметим, что одно из четырех указанных уравнений может быть получено как следствие из других. Так, например, дифференцируя уравнение квазинейтральности (4.11), мы имеем

$$\frac{dp}{dt} = \frac{dn}{dt} + N_i \frac{df}{dt}.$$

Подставляя сюда для каждого из слагаемых их выражения (4.8) — (4.10), мы получаем тождество. Остающиеся же три независимых уравнения определяют три искомые величины $n(t)$, $p(t)$ и $f(t)$.

Так как уравнения (4.8) и (4.9) нелинейны, то в общем случае закон изменения концентраций n и p не выражается экспонентой. Однако в специальных случаях существуют постоянные характерные времена процесса. Ниже приведены некоторые примеры.

7-8. ТЕМПЕРАТУРНАЯ ЗАВИСИМОСТЬ ВРЕМЕНИ ЖИЗНИ НОСИТЕЛЕЙ ЗАРЯДА ПРИ РЕКОМБИНАЦИИ ЧЕРЕЗ РЕКОМБИНАЦИОННЫЕ ЛОВУШКИ

Рассмотрим теперь, как меняется время жизни носителей заряда при рекомбинации через рекомбинационные ловушки в полупроводнике заданного состава при изменении температуры исследования. Этот анализ проведем на примере полупроводника электронного типа. Пусть в таком примесном полупроводнике при повышении температуры наблюдаются все три области изменения концентрации носителей заряда, как это изображено на рис. 7-9, т. е. имеется область примесной проводимости — область вымораживания (участок А), область истощения, т. е. область полной ионизации примеси, характеризующаяся постоянством концентрации носителей заряда (участок Б), и область собственной проводимости (участок В).

В области А уровень Ферми, температурный ход которого изображен на рис. 7-9, б, лежит между уровнями донорной примеси и дном зоны проводимости. С повышением температуры уровень Ферми понижается до положения уровня донорной примеси. Поскольку в области А уровень Ферми для всех температур лежит выше уровня энергии рекомбинационной ловушки, концентрация электронов подчиняется неравенству $n_0 \gg n_1$, $n_0 \gg p_0$ и $n_0 \gg p_1$. Следовательно, в этой области в соответствии с выражением (7-86) время жизни постоянно и равно τ_{p0} . Это участок 1' на рис. 7-9, в, на котором представлена температурная зависимость времени жизни.

По мере дальнейшего повышения температуры при переходе к области полной ионизации донорной примеси уровень Ферми понижается и достигает уровня рекомбинационной ловушки. На этом участке концентрация электронов n_0 постоянна, а концентрация $n_1 = N_c e^{-(E_c - E_l)/kT}$ растет с температурой экспоненциально, но она по-прежнему значительно меньше равновесной концентрации n_0 . Кроме того, n_0 также больше концентраций дырок p_0 и p_1 . На этом участке согласно выражению (7-85) время жизни также по-

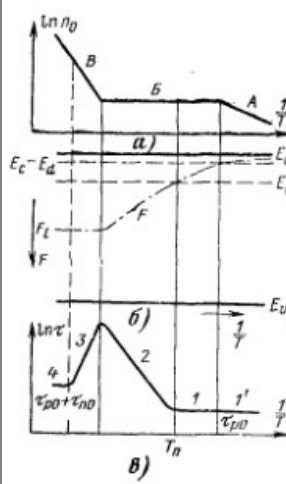


рис. 7-9. Зависимость равновесной концентрации (а), уровня Ферми (б) и времени жизни (в) от температуры для донорного полупроводника.

стоянно и равно τ_{p0} (участок 1 на рис. 7-9, в).

При дальнейшем повышении температуры уровень Ферми понижается и находится ниже энергетического уровня рекомбинационной ловушки (участок 2 на рис. 7-9, б). Здесь по-прежнему $n_0 = \text{const}$, а n_1 значительно увеличивается, поэтому из выражения (7-90) будем иметь:

$$\tau = \tau_{p0} \frac{n_1}{n_0} = \tau_{p0} \frac{N_c}{n_0} e^{-(E_c - E_l)/kT} = \frac{\tau_{p0}}{n_0} 2 \left(\frac{2\pi m_n^* k}{h^2} \right)^{3/2} T^{3/2} e^{-(E_c - E_l)/kT}, \quad (7-93)$$

т. е. время жизни увеличивается с ростом температуры. При этом оно увеличивается вплоть до температуры, при которой наступает собственная проводимость (участок 2 на рис. 7-9, в). Рост τ в этой области температур происходит потому, что с увеличением температуры имеет место интенсивная эмиссия электронов с уровня рекомбинационных ловушек. В этом случае уменьшается заполнение ловушек электронами, а следовательно, увеличивается время жизни неосновных носителей — дырок, которое и определяет измеряемое время τ .

В области собственной проводимости с ростом температуры, т. е. положение уровня Ферми практически остается постоянным (рис. 7-9, б), концентрация носителей заряда в зонах резко увеличивается, поэтому увеличивается заполнение рекомбинационных ловушек носителями заряда и время жизни уменьшается. В этой области

$$n_0 \approx p_0 \approx n_i = (N_c N_v)^{1/2} e^{-E_g/2kT}. \quad (7-94)$$

В соответствии с этим соотношение (7-85) запишется в виде

$$\begin{aligned} \tau &= \tau_{p0} \frac{n_i + n_1}{2n_i} + \tau_{n0} \frac{n_i + p_1}{2n_i} = \\ &= \frac{\tau_{p0}}{2} \left[1 + \left(\frac{N_c}{N_v} \right)^{1/2} e^{\frac{E_g + E_l - E_c}{2kT}} \right] + \frac{\tau_{n0}}{2} \left[1 + \left(\frac{N_v}{N_c} \right)^{1/2} e^{\frac{E_g - E_l - E_v}{2kT}} \right] = \\ &= \frac{\tau_{p0}}{2} \left[1 + \left(\frac{N_c}{N_v} \right)^{1/2} e^{\frac{E_c + E_v - E_l}{2kT}} \right] + \frac{\tau_{n0}}{2} \left[1 + \left(\frac{N_v}{N_c} \right)^{1/2} e^{\frac{E_c + E_v - E_l}{2kT}} \right]. \end{aligned} \quad (7-95)$$

Если уровень рекомбинационных ловушек расположен достаточно далеко от середины запрещенной зоны, то в области начала собственной проводимости можно выделить участок, для которого выполняется неравенство $\left(\frac{E_c + E_v}{2} - E_l \right) \gg kT$. Для этого случая уравнение (7-95) приобретает вид:

$$\tau \approx \frac{\tau_{p0}}{2} \left(\frac{N_v}{N_c} \right)^{1/2} e^{\frac{E_c + E_v - E_l}{2kT}}. \quad (7-96)$$

Соответствующий участок на кривой $\ln \tau \left(\frac{1}{T} \right)$ характеризуется уменьшением времени жизни (участок 3 на рис. 7-9, в).

При дальнейшем повышении температуры, когда в области собственной проводимости будет иметь место неравенство $\left(\frac{E_c + E_v}{2} - E_l \right) \ll kT$, уравнение (7-95) можно записать приближенно следующим образом:

$$\tau = \frac{\tau_{p0}}{2} \left(1 + \sqrt{\frac{N_c}{N_v}} \right) + \frac{\tau_{n0}}{2} \left(1 + \sqrt{\frac{N_v}{N_c}} \right). \quad (7-97)$$

В этом случае время жизни стремится к постоянной величине. И если принять $N_c = N_v$, это выражение упрощается:

$$\tau = \tau_{p0} + \tau_{n0} = \frac{\gamma_n + \gamma_p}{\gamma_n \gamma_p N_i}. \quad (7-98)$$

Таким образом, когда в области собственной проводимости с увеличением температуры концентрация свободных носителей заряда становится очень высокой, время жизни определяется только параметрами рекомбинационных ловушек и не зависит от температуры (участок 4 на рис. 7-9, в).

Зная концентрацию электронов при насыщении, можно оценить величину E_l по температуре перехода T_{12} от области 1 к области 2 (рис. 7-9, в). При условии $T = T_{12}$, а $F = E_l$ находим:

$$n_0 = N_c e^{-(E_c - E_l)/kT_{12}}, \quad (7-99)$$

откуда

$$E_c - E_l = -kT_{12} \ln \frac{N_c}{n_0}. \quad (7-100)$$

Экстраполируя прямую область 2 до пересечения с осью ординат, определяем в соответствии с равенством (7-93) значение $\ln(\tau_{p0} N_c / n_0)$, а зная τ_{p0} из участка 1, находим $\ln(N_c / n_0)$ и подставляем в формулу (7-100). Следовательно, по снятой экспериментальной зависимости $\ln \tau \left(\frac{1}{T} \right)$, подобной изображенной на рис. 7-9, в, можно рассчитать глубину залегания уровня рекомбинационных ловушек E_l .

7.9. ЛОВУШКИ ЗАХВАТА И РЕКОМБИНАЦИОННЫЕ ЛОВУШКИ

При анализе процессов рекомбинации носителей заряда через рекомбинационные ловушки мы учитывали захват электронов и дырок с ловушек в соответствующие зоны. При этом явление теплового возбуждения захваченных носителей заряда с уровня ловушки приводит к увеличению времени жизни, а следовательно, к уменьшению скорости рекомбинации. Очевидно, соотношение интенсивности этих процессов зависит от положения энергетического уровня ловушек.

На рис. 7-10 изображена энергетическая схема полупроводника, содержащего локальные центры, характеризующиеся различными уровнями в запрещенной зоне. Если мы имеем состояние, отличное от термодинамического равновесия, значения концентраций свободных носителей заряда определяются положением квазиуровней Ферми для электронов F_n и дырок F_p . Если энергетический уровень ловушек расположен близко к дну зоны проводимости, вероятность тепловой ионизации носителей заряда велика. Такие центры находятся в состоянии непрерывного обмена электронами с зоной проводимости и не вносят существенного вклада в процессы рекомбинации. Это ловушки захвата электронов. Вблизи верхнего края валентной зоны располагаются уровни ловушек захвата дырок.

У рекомбинационных ловушек, которые осуществляют захват электронов и дырок, в результате чего происходит их рекомбинация, тепловой заброс носителей заряда в соответствующие зоны маловероятен, поэтому рекомбинационные ловушки должны характеризоваться энергетическими уровнями, расположенными достаточно глубоко под дном зоны проводимости.

Охарактеризовать соотношение процессов захвата и рекомбинации количественно можно, введя коэффициент k , выражающий собой вероятность захвата носителя заряда противоположного знака.

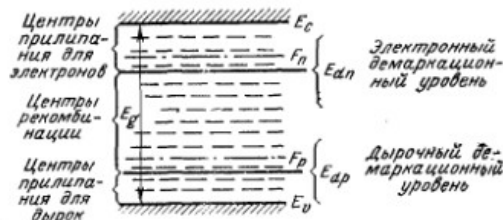


Рис. 7-10. Полупроводник, содержащий ловушки захвата и рекомбинационные ловушки.

$$k_p = \frac{\gamma_n n}{\gamma_p p_1} \quad (7-104)$$

Из условия $k_p = 1$ найдем положение дырочного демаркационного уровня:

$$\gamma_n N_c e^{-(E_c - F_n)/kT} = \gamma_p N_v e^{(E_v - E_{dp})/kT} \quad (7-105)$$

или

$$E_{dp} = E_c + E_v - F_n - kT \ln \frac{\gamma_n N_c}{\gamma_p N_v} \quad (7-106)$$

Таким образом, если энергетический уровень ловушек лежит выше электронного демаркационного уровня, они являются ловушками захвата для электронов. Ловушки, которые имеют уровни энергии ниже дырочного демаркационного уровня, являются ловушками захвата для дырок. Между ними располагаются энергетические уровни рекомбинационных ловушек (рис. 7-10). Необходимо отметить, что такая классификация энергетических уровней ловушек является весьма условной и схематичной. При ее проведении, в частности, считалось, что отношение коэффициентов захвата γ_n и γ_p изменяется сравнительно слабо при перемещении уровня ловушек. Ловушки захвата могут также вносить некоторый вклад и в процессы рекомбинации, который становится особенно заметным, если в полупроводнике отсутствуют дефекты, дающие энергетические уровни в средней части запрещенной зоны.

Он равен отношению вероятности захвата к вероятности теплового заброса. Применительно к электронам зоны проводимости k_n будет равен отношению вероятности захвата отрицательно заряженной ловушкой дырки к вероятности теплового заброса электрона с ловушки в зону проводимости. Как следует из выражения (7-70), скорость захвата ловушкой дырок определяется величиной $\gamma_p N_v f_p$. Интенсивность процесса теплового возбуждения электронов с ловушек согласно формуле (7-69) равна $\gamma_n N_c f_n$. Здесь γ_n , γ_p — вероятность захвата электрона и дырки соответственно; f_n — неравновесная функция распределения для электронов, находящихся на ловушках; p — концентрация неравновесных дырок в валентной зоне; n_1 — равновесная концентрация электронов в зоне проводимости в том случае, когда уровень Ферми совпадает с уровнем ловушки.

Учитывая это, для коэффициента k_n получаем следующее выражение:

$$k_n = \frac{\gamma_p p}{\gamma_n n_1} = \frac{\gamma_p p}{\gamma_n N_c e^{-(E_c - E_l)/kT}} \quad (7-101)$$

Ловушки, для которых вероятность рекомбинации больше вероятности теплового освобождения, т. е. $k_n > 1$, будут являться рекомбинационными ловушками, а ловушки с $k_n < 1$ относятся к ловушкам захвата. Энергетический уровень ловушки, который характеризуется значением $k_n = 1$, т. е. для которого вероятность рекомбинации и теплового освобождения одинакова, называется электронным демаркационным уровнем E_{dn} . Его положение можно найти из условия

$$\gamma_p N_c e^{(E_v - F_p)/kT} = \gamma_n N_c e^{-(E_c - E_{dn})/kT} \quad (7-102)$$

Отсюда будем иметь:

$$E_{dn} = E_c + E_v - F_p - kT \ln \frac{\gamma_n N_c}{\gamma_p N_v} \quad (7-103)$$

Из полученного соотношения следует, что положение электронного демаркационного уровня E_{dn} зависит от ширины запрещенной зоны, положения квазиуровня Ферми дырок, эффективной массы электронов и дырок, параметров ловушек (γ_n , γ_p) и температуры. С увеличением уровня возбуждения, приводящего к росту концентрации дырок, уменьшается величина $F_p - E_v$ и электронный демаркационный уровень приближается к дну зоны проводимости. В этом случае будет осуществляться переход ловушек захвата в рекомбинационные ловушки. С ростом же температуры демаркационный уровень сдвигается к середине запрещенной зоны и рекомбинационные ловушки будут переходить в ловушки захвата.

Для рекомбинационных ловушек и ловушек захвата, уровни которых расположены вблизи верхнего края валентной зоны, можно ввести коэффициент

10.4. ПОВЕРХНОСТНАЯ РЕКОМБИНАЦИЯ

Из большого числа состояний, возникающих на поверхности полупроводника, некоторая часть может являться эффективными рекомбинационными ловушками. Это может привести к тому, что темп рекомбинации на поверхности будет значительно выше, нежели в объеме полупроводника. Рекомбинационно активными ловушками на поверхности могут быть лишь быстрые состояния, поскольку время перехода на медленные состояния настолько велико, что не может быть обеспечен темп рекомбинации, по порядку величины сравнимый с объемной рекомбинацией. Влияние поверхностной рекомбинации на время жизни носителей заряда становится существенным тогда, когда число электронно-дырочных пар, рекомбинирующих в единицу времени на поверхности, составляет заметную долю от общего числа носителей заряда, рекомбинирующих во всем объеме полупроводника. Можно поэтому ожидать, что влияние поверхностной рекомбинации будет особенно сильно в случае тонких пластин и нитевидных образцов, имеющих большое отношение поверхности к объему.

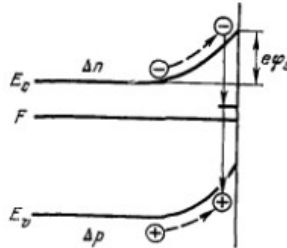


Рис. 10-13. Рекомбинация через поверхностные уровни.

Для характеристики процесса рекомбинации на поверхности вводят величину s , называемую скоростью поверхностной рекомбинации.

Пусть на поверхности электронного полупроводника имеются дефекты с концентрацией N_s , характеризующиеся энергетическим уровнем E_s (рис. 10-13), через который осуществляется процесс рекомбинации по модели Шокли—Рида, рассмотренной в § 7-7.

В условиях термодинамического равновесия поток дырок к поверхности полупроводника в соответствии с (9-85) равен $\frac{1}{4} v_0 p_0$,

где v_0 — тепловая скорость дырок; p_0 — равновесная концентрация дырок. Часть этих дырок будет захвачена на уровне E_s , а часть дырок, равная $\frac{1}{4} r v_0 p_0$, отразится от поверхности, где r — сред-

няя скорость отражения дырки от поверхности полупроводника. Количество дырок, захваченных поверхностными уровнями в единицу времени, в условиях равновесия равно числу дырок q_{s0} , генерируемых за это же время с поверхностных уровней в валентную зону. Но при термодинамическом равновесии поток дырок, подходящих к поверхности полупроводника, равен потоку дырок, движущихся от поверхности. Обратный поток включает как отраженные от поверхности дырки, так и дырки, освобожденные с поверхностных уровней. Поэтому

$$\frac{1}{4} v_0 p_0 = \frac{1}{4} r v_0 p_0 + q_{s0}, \quad (10-36)$$

откуда скорость ухода дырок с поверхностных уровней в состоянии равновесия

$$q_{s0} = \frac{1}{4} (1-r) v_0 p_0. \quad (10-37)$$

При отклонении от равновесия поток дырок к поверхности $q_s = \frac{1}{4} (1-r) v_0 p$ и поток дырок от поверхности уже не компенсируют друг друга. Можно ожидать, что при малых отклонениях от равновесия количество дырок, генерируемых с поверхностного уровня, не зависит от величины отклонения, и следовательно $q_{s0} = \frac{1}{4} (1-r) v_0 p_0$. Таким образом, будет существовать результирующий поток дырок к поверхности, равный:

$$\Delta q_s = q_s - q_{s0} = \frac{1}{4} (1-r) v_0 (p - p_0) = \frac{1}{4} (1-r) v_0 \Delta p, \quad (10-38)$$

или

$$\Delta q_s = s \Delta p, \quad (10-39)$$

где

$$s = \frac{1}{4} (1-r) v_0. \quad (10-40)$$

Величина s , являющаяся коэффициентом пропорциональности между Δq_s и избыточной концентрацией дырок Δp и имеющая размерность скорости, носит название скорости поверхностной рекомбинации.

При таком способе введения понятия s результирующий поток дырок к поверхности полупроводника определяется различием между потоком дырок, захваченных на уровне E_s , и потоком дырок, термически выброшенных с этих уровней в валентную зону, т. е. зависит только от «пропускной способности» поверхностного рекомбинационного уровня. Тогда для скорости изменения концентрации дырок в результате рекомбинации через поверхностные уровни может быть записано соотношение, аналогичное (7-80) в теории Шокли—Рида:

$$-\left(\frac{dp}{dt}\right)_r = R_p = \frac{\gamma_{ns} \gamma_{ps} (n_0 p_0 - n_1 p_1)}{\gamma_{ns} (n_s + n_1) + \gamma_{ps} (p_s + p_1)}, \quad (10-41)$$

где γ_{ns} и γ_{ps} — всраятность захвата поверхностным уровнем электронов и дырок из соответствующих зон в течение одной секунды, имеющая размерность $\text{см}^2/\text{с}$; n_1 и p_1 — неравновесные концентрации электронов и дырок на поверхности полупроводника; n_0 и p_0 — равновесные концентрации электронов и дырок, когда уровень Ферми совпадает с уровнем рекомбинационной ловушки E_s , поэтому $n_1 p_1 = n_0 p_0 = n_i^2$.

Поскольку в стационарных условиях $-\left(\frac{dp}{dt}\right)_r = -\left(\frac{dn}{dt}\right)_r$ или $R_p = R_n = R$, то имеют место равные по величине потоки дырок и электронов к поверхности, в противном случае на поверхности полупроводника имело бы место накопление носителей заряда одного знака и, следовательно, состояние не было бы стационарным. Результирующий поток электронов и дырок к поверхности обусловлен градиентом избыточных носителей заряда, возникающим в приповерхностной области полупроводника в результате различия в темпах рекомбинации на поверхности и в объеме полупроводника. И если скорость поверхностной рекомбинации невелика, то число пар носителей заряда, рекомбинирующих на поверхности полупроводника, невелико по сравнению с потоками к поверхности и от поверхности. А эти потоки связаны с диффузионными и дрейфовыми токами в области пространственного заряда, поэтому при малой скорости поверхностной рекомбинации можно считать, что диффузионные и дрейфовые токи практически уравновешивают друг друга. В этом случае концентрации носителей заряда будут определяться распределением Больцмана. Если величина изгиба зоны у поверхности $e\phi_s$, а n и p есть концентрации неравновесных электронов и дырок в объеме полупроводника, то их поверхностные значения равны:

$$n_s = n e^{e\phi_s/kT} = n e^{Y_s}; \quad p_s = p e^{-e\phi_s/kT} = p e^{-Y_s}, \quad (10-42)$$

а в условиях термодинамического равновесия

$$n_{s0} = n_0 e^{Y_s}; \quad p_{s0} = p_0 e^{-Y_s}. \quad (10-43)$$

С учетом (10-42) равенство (10-41) запишется в виде

$$R = \frac{\gamma_{ns} \gamma_{ps} (n p - n_0 p_0)}{\gamma_{ns} (n_s + n_1) + \gamma_{ps} (p_s + p_1)}. \quad (10-44)$$

Имея в виду, что $p = p_0 + \Delta p$ и $n = n_0 + \Delta n$, а $\Delta n = \Delta p$ в глубине полупроводника, при условии малого уровня возбуждения, когда $\Delta n, \Delta p \ll n_0, p_0$,

$$R = \frac{\gamma_{ns} \gamma_{ps} (n_0 + p_0) \Delta p}{\gamma_{ns} (n_{s0} + n_1) + \gamma_{ps} (p_{s0} + p_1)} = \Delta q_s = s \Delta p. \quad (10-45)$$

Используя (10-43), соотношение для скорости поверхностной рекомбинации в зависимости от величины поверхностного электростатического потенциала Y_s приобретает вид:

$$s = \frac{\gamma_{ns} \gamma_{ps} (n_0 + p_0)}{\gamma_{ns} (n_0 e^{Y_s} + n_1) + \gamma_{ps} (p_0 e^{-Y_s} + p_1)}. \quad (10-46)$$

Зависимость $s(Y_s)$ имеет максимум при значениях $Y_{\text{макс}}$, которое можно найти из условия

$$Y_{\text{макс}} = \frac{1}{2} \ln \frac{\gamma_{ns}^2 P_0}{\gamma_{ps} P_0} \quad (10-47)$$

Таким образом, скорость поверхностной рекомбинации s можно определить как число актов рекомбинации на 1 см^2 поверхности полупроводника в 1 с ; отнесенных к избыточной концентрации носителей заряда на границе области пространственного заряда.

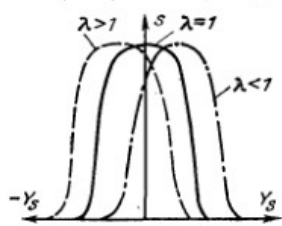


Рис. 10-14. Зависимость скорости поверхностной рекомбинации от безразмерного поверхностного электростатического потенциала.

Вид зависимости $s(Y_s)$ в соответствии с (10-46) представлен на рис. 10-14. Из (10-46) очевидно, что если из эксперимента известно значение электростатического потенциала, соответствующее максимуму скорости поверхностной рекомбинации, то можно вычислить отношение вероятностей захвата электрона и дырки рекомбинационным уровнем или, поскольку $\gamma_{ns,ps} = \sigma_{n,p} v_0 N_s$, можно определить отношение сечений захвата электрона и дырки σ_n/σ_p . Качественное объяснение «колоколообразной» зависимости $s(Y_s)$

состоит в том, что при изменении Y_s меняется положение рекомбинационного уровня на поверхности полупроводника относительно дна зоны проводимости и потолка валентной зоны и, как следствие этого, его эффективность как рекомбинационной ловушки.

Виды люминесценции

Излучение является процессом, обратным поглощению. Электрон, занимающий состояние с энергией, более высокой чем в равновесных условиях, совершает переход в незаполненное низкоэнергетическое состояние; при этом вся или большая часть разницы энергий этих состояний может выделиться в виде электромагнитного излучения. Скорость такого излучения равна произведению плотности n_u носителей в верхнем состоянии на плотность n_l пустых состояний с более низкой энергией и на вероятность P_{ul} излучательного перехода при плотности 1 носитель/см³ в верхнем состоянии и при плотности 1 вакансии/см³ в нижнем состоянии:

$$R = n_u n_l P_{ul} \quad (6.1)$$

Это несколько напоминает выражение (3.1), которое мы писали для процесса поглощения; однако там поглощение описывалось нами с помощью понятия среднего свободного пробега фотона, тогда как излучение мы описываем теперь через скорость генерации фотонов в единице объема.

Большинство переходов, которые мы рассмотрели как примеры механизмов поглощения, может идти и в обратном направлении, создавая *характерное излучение*. Однако существует важное различие между той информацией, которую можно получить при излучении поглощения, и той, которую можно получить при исследовании излучения в полупроводнике: в процессы поглощения могут вовлекаться все состояния в полупроводнике по обеим сторонам от уровня Ферми, что приводит к широкому спектру, в то время как излучательные переходы происходят между узкой полосой состояний, в которой находятся термализованные электроны, и узкой полосой пустых состояний, содержащих термализованные дырки, и, следовательно, создают узкий спектр.

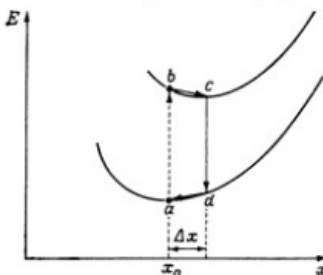
Главным условием для возникновения излучения является то, что система не должна быть в равновесии. Для отклонения от равновесия требуется одна из форм возбуждения. Процесс излучения света называют люминесценцией. Возбуждение посредством электрического тока (инжекция или пробой) приводит к электролюминесценции. Оптическое возбуждение (при поглощении) создает фотолюминесценцию, а возбуждение электронным пучком — катодолуминесценцию. Триболюминесценция обусловлена механическим возбуждением. Однако термолуминесценция — не просто следствие теплового возбуждения (в этом случае имело бы место температурное свечение («инканденсенция»)); для термолуминесценции требуется возбуждение при низкой температуре, чтобы выморозить носители на ловушки, из которых они могут быть потом высвобождены термически. При хемолуминесценции свет излучается вследствие химической реакции, но никто пока не сообщал о подобном явлении в полупроводниках.

Флюоресценция — это люминесценция, которая происходит только во время возбуждения. Фосфоресценция — это люминесценция, которая продолжается некоторое время после окончания возбуждения.

Стоківський зсув

В этом параграфе мы вводим представление о конфигурационной диаграмме, так как она полезна при объяснении процессов испускания фононов, сопровождающих многие излучательные переходы.

Конфигурационная диаграмма (фиг. 6.3) изображает энергию основного и первого возбужденного состояния атома как функцию его положения, причем этот атом может быть или примесным, или собственным атомом решетки. Отметим, что в основном состоянии равновесному положению атома соответствует минимум энергии. Эта минимальная энергия есть равновес-



Фиг. 6.3. Конфигурационная диаграмма.

ная энергия электрона в основном состоянии, когда атом неподвижен в своем равновесном положении в неподвижной решетке (такая ситуация возможна при очень низкой температуре). Однако возбужденное состояние может иметь минимальную энергию при несколько другом положении атома. Следовательно, при возбуждении электрона из основного состояния вся система частично релаксирует в состояние c . Это означает смещение атома на Δx . В этом процессе возбужденный электрон теряет некоторую энергию, которая рассеивается в виде атомных смещений, т. е. в виде фононов.

При возвращении электрона в основное состояние (точка d) необходимо новое смещение атома, для того чтобы система пришла к своей наиболее низкой энергии в точке a . Это смещение $d \rightarrow a$ также происходит путем испускания фонона.

Хотя в этом процессе могут принимать участие все или несколько типов фононов, обычно наиболее вероятен LO-фонон, так как

именно он создает самое сильное поляризационное поле, т. е. наибольшее изменение потенциала на единицу смещения.

Энергия, излучаемая при переходе $c \rightarrow d$, меньше чем энергия, поглощаемая при переходе $a \rightarrow b$. Разницу этих двух энергий называют «сдвигом Франка — Кондона». Вообще же любое уменьшение оптической энергии называют стоксовским смещением.

Таким образом, сдвиг Франка — Кондона — это стоксовское смещение, обусловленное сдвигом атомов после оптического возбуждения.

§ 1. СООТНОШЕНИЕ ВАН РУСБРЕКА — ШОКЛИ [1]

В равновесии скорость оптической генерации электронно-дырочных пар равна скорости их излучательной рекомбинации — такова простая формулировка соотношения ван Русбрека — Шокли.

Если для различных частот фотонов ν составить детальный баланс этих процессов, то можно написать, что скорость излучения на частоте ν , лежащей в интервале $d\nu$, равна

$$R(\nu) d\nu = P(\nu) \rho(\nu) d\nu, \quad (6.2)$$

где $P(\nu)$ — вероятность поглощения фотона с энергией $h\nu$ в единицу времени, а $\rho(\nu) d\nu$ — плотность фотонов с частотой ν в интервале $d\nu$. Последнюю величину можно получить из закона излучения Планка (предполагая для простоты, что коэффициент преломления n не зависит от ν):

$$\rho(\nu) d\nu = \frac{8\pi\nu^2 n^3}{c^3} \frac{1}{\exp(h\nu/kT) - 1} d\nu. \quad (6.3)$$

Вероятность поглощения связана со средним временем жизни $\tau(\nu)$ фотона в полупроводнике:

$$P(\nu) = \frac{1}{\tau(\nu)}.$$

Среднее время жизни можно рассчитать из среднего свободного пробега $1/\alpha(\nu)$ для фотона, движущегося со скоростью $v = c/n$ (опять-таки предполагая коэффициент преломления постоянным):

$$\tau(\nu) = \frac{1}{\alpha(\nu) v}.$$

Следовательно,

$$P(\nu) = \alpha(\nu) v = \alpha(\nu) \frac{c}{n}. \quad (6.4)$$

Подставляя (6.3) и (6.4) в (6.2) получаем

$$R(\nu) d\nu = \frac{\alpha(\nu) 8\pi\nu^2 n^2}{c^2 [\exp(h\nu/kT) - 1]} d\nu. \quad (6.5)$$

Равенство (6.5) представляет собой фундаментальное соотношение между наблюдаемым спектром поглощения и ожидаемым спектром излучения. Скорость излучения (6.5) может быть выражена через коэффициент экстинкции $k(\nu)$. Напомним, что, согласно (4.7),

$$\alpha(\nu) = \frac{4\pi\nu k(\nu)}{c}, \quad (6.6)$$

поэтому

$$R(\nu) d\nu = \frac{32\pi^2 k(\nu) n^2 \nu^3}{c^3 [\exp(h\nu/kT) - 1]} d\nu. \quad (6.7)$$

Формула состоит из двух сомножителей, n^2 и $\alpha(\nu)$, характеризующих данный полупроводник, и множителя, не зависящего от химического состава вещества

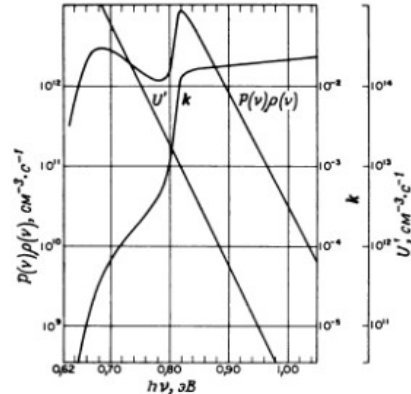
$$U = \frac{8\pi}{c^2} \left(\frac{kT}{h}\right)^3 \frac{u^3}{e^u - 1}.$$

Если же использовать k вместо α [2], то множителем, не зависящим от природы вещества, будет

$$U' = \frac{32\pi^2}{c^3} \left(\frac{kT}{h}\right)^4 \frac{u^3}{e^u - 1} = 1,785 \cdot 10^{22} \left(\frac{T}{300}\right)^4 \frac{u^3}{e^u - 1}.$$

Следовательно, в принципе возможно для полупроводника с известным коэффициентом преломления n трансформировать спектр поглощения в ожидаемый спектр излучения (см., например, фиг. 6.1).

Равенства (6.5) и (6.8) устанавливают фундаментальную связь между скоростью рекомбинации и коэффициентом поглощения. Хотя вышеприведенная формулировка была первоначально выведена для переходов зона — зона, она справедлива и для переходов



Ф и г. 6.1. Зависимости $k = ca/4\pi\nu$, U' и $P(\nu)\rho(\nu)$ от энергии для германия при 300 K [2].

между любыми совокупностями состояний. Пока мы рассматривали случаи термического равновесия; теперь посмотрим, что произойдет при отклонении от термического равновесия.

Мы сохраним символы n и p для концентрации электронов и дырок соответственно, но будем подразумевать, что p относится к концентрации любого сорта незаполненных состояний ниже уровня Ферми, т. е. это могут быть и дырки в валентной зоне, и ионизированные доноры или нейтральные акцепторы, и дырки, относящиеся к экситону.

Пусть n_i — собственная концентрация носителей. Общая скорость излучательной рекомбинации R_c пропорциональна концентрации электронов n и концентрации пустых состояний p :

Равенство (6.9) показывает, что время жизни носителей уменьшается при возрастании концентрации носителей, а также то, что общая скорость излучательной рекомбинации становится равной R , когда произведение np приближается к собственному равновесному значению n_i^2 .

Выразим концентрации носителей через отклонения от равновесных значений: $n = n_0 + \Delta n$ и $p = p_0 + \Delta p$, тогда скорость рекомбинации примет вид

$$R + \Delta R = \frac{(n_0 + \Delta n)(p_0 + \Delta p)}{n_0 p_0} R = \frac{n_0 p_0 + p_0 \Delta n + n_0 \Delta p + \Delta n \Delta p}{n_0 p_0} R$$

или в пренебрежении малым членом $\Delta n \Delta p$

$$\frac{\Delta R}{R} = \frac{\Delta n}{n_0} + \frac{\Delta p}{p_0}. \quad (6.10)$$

Теперь мы можем найти излучательное время жизни неравновесных носителей, предполагая $\Delta n = \Delta p$:

$$\tau = \frac{\Delta n}{\Delta R} = \frac{1}{R} \frac{n_0 p_0}{n_0 + p_0}. \quad (6.11)$$

Следовательно, для собственного вещества, когда $n_0 = p_0 = n_i$,

$$\tau = \frac{n_i}{2R}. \quad (6.12)$$

Если обозначить через B вероятность излучательной рекомбинации (она эквивалентна P_{ul} в формуле (6.1), где рассматривались только дискретные состояния, а не набор состояний), то

$$B = \frac{R}{n_i^2}. \quad (6.13)$$

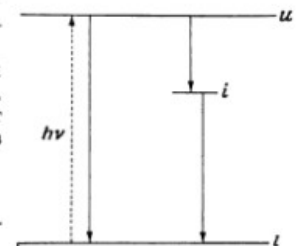
Расчитанные значения τ и B для излучательных переходов «зона проводимости — валентная зона» приведены в табл. 6.1.

Эффективность вибр. фотолумин.(квантовый выход)

Переходы из верхнего в более низкоэнергетическое состояние могут происходить также через одно или несколько промежуточных состояний. Промежуточные переходы могут быть излучательными или безизлучательными. Тем не менее они существенно влияют на эффективность излучательных переходов из верхнего в нижнее состояние.

Рассмотрим случай одного промежуточного состояния i (фиг. 6.2). Расчеты ван Русбрека — Шокли для вероятности излучательных переходов $u \rightarrow l$ основаны на знании измеренного при этой же энергии $h\nu$ коэффициента поглощения для обратного процесса $l \rightarrow u$.

Ясно, что в модели дискретных состояний, изображенной на фиг. 6.2, промежуточный уровень не участвует в процессе поглощения фотонов



Ф и г. 6.2. Конкурирующие рекомбинационные процессы.

с энергией $h\nu$. Для рекомбинационных же процессов возможны два пути, и, следовательно, имеют место два конкурирующих процесса: излучательный переход с испусканием фотона $h\nu$ и переход $u \rightarrow i \rightarrow l$, который (если он носит излучательный характер) дает фотоны с энергией, меньшей $h\nu$. Часто переходы $u \rightarrow i \rightarrow l$ доминируют и делают слабыми излучательные переходы с энергией $h\nu$. Тем не менее излучательные переходы $u \rightarrow l$ по-прежнему описываются формулой (6.5). Однако число носителей, принимающих участие в этом процессе, уменьшено влиянием конкурирующего процесса, который может быть охарактеризован временем рекомбинации τ' . Следовательно, эффективное время рекомбинации определяется из равенства

$$\frac{1}{\tau_{\text{эфф}}} = \frac{1}{\tau} + \frac{1}{\tau'} \quad (6.14)$$

Комбинируя (6.14), (6.12) и (6.9), получаем общую скорость рекомбинации

$$R_T = \frac{1}{\tau_{\text{эфф}}} \frac{n_p}{2n_i}$$

в то время как скорость излучательной рекомбинации равна. Следовательно, эффективность излучения равна

$$\eta = \frac{R}{R_T} = \frac{1}{1 + \tau/\tau'} \quad (6.15)$$

Статистика рекомбинации через промежуточное состояние с учетом концентрации основных и неосновных носителей рассмотрена Шокли и Ридом [5], а также Холлом [3].

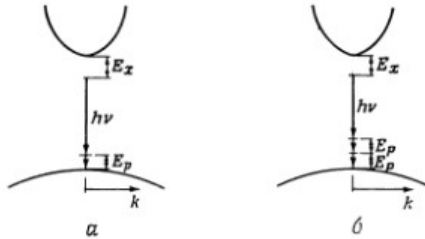
в n^2 раз меньше, чем энергия ионизации основного состояния E_{x1} , соответствующего $n = 1$. Следовательно, излучение свободного экситона могло бы состоять из серии узких линий, начинающихся при $E_g - E_{x1}$ и продолжающихся при энергиях $E_g - (1/n^2)E_{x1}$. Однако интенсивность пиков более высокого порядка быстро уменьшается [6] (а именно как n^{-3}), и в присутствии основного перехода трудно наблюдать другие излучательные переходы. В GaAs обнаружены свободные экситоны с $n = 1$ и $n = 2$ [7].

В полупроводнике с непрямой запрещенной зоной сохранение импульса требует, чтобы переход был дополнен испусканием фонона (фиг. 6.4, б). Тогда энергия излученного фотона равна

$$h\nu = E_g - E_x - E_p,$$

где E_p — энергия участвующего в переходе фонона.

Отметим, что прямые переходы также могут происходить с испусканием одного или более фононов (фиг. 6.5), но при этом вероятность перехода уменьшается. Действительно, рассмотрение



Фиг. 6.5. Прямая экситонная рекомбинация.

а — с испусканием одного оптического фонона; б — с испусканием двух оптических фононов.

конфигурационной диаграммы показывает, что даже прямые переходы могут, вообще говоря, приводить к испусканию фононов, причем наиболее вероятен оптический фонон (оптические фононы возможны и при $k = 0$). Таким образом, узкий спектр излучения экситона может повториться при несколько меньших энергиях фотонов, таких, что

$$h\nu = E_g - E_x - mE_p,$$

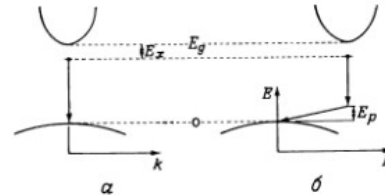
где m — число оптических фононов, излученных при переходе. Очевидно, что чем больше m , тем ниже вероятность перехода и слабее соответствующая линия излучения.

Следует заметить, что, хотя переходы с испусканием фонона менее вероятны, чем прямая рекомбинация, образовавшийся при этом фотон имеет больший шанс выйти из кристалла, так как он обладает меньшей энергией, чем фотон от прямой рекомбинации, и находится в области спектра, где полупроводник более прозрачен. Однако реабсорбция экситонного излучения не обязательно означает потерю энергии, поскольку в этом процессе создается новый экситон, чем обеспечивается новая возможность для излучения.

1. Экситонная рекомбинация

а. Свободный экситон

Если вещество достаточно чистое, то электроны и дырки спариваются в экситон, который затем рекомбинирует, излучая узкую



Фиг. 6.4. Экситонная рекомбинация: а — прямой; б — непрямой.

спектральную линию. В полупроводнике с прямой запрещенной зоной, где сохранение импульса возможно при простом излучательном переходе (фиг. 6.4, а), энергия излученного фотона равна

$$h\nu = E_g - E_x. \quad (6.16)$$

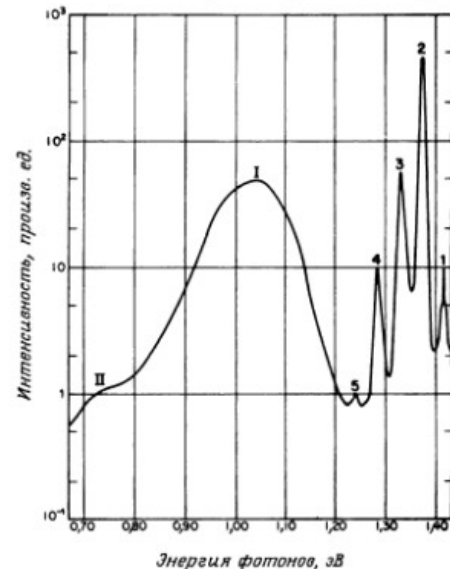
В п. 1 § 4 гл. 1 мы видели, что экситон может иметь последовательность возбужденных состояний, энергии ионизации которых

в. Связанный экситон

В присутствии примесей могут образовываться связанные экситоны. Когда они рекомбинируют, излучение состоит из узких линий с энергиями, меньшими, чем при рекомбинации свободного экситона. В GaAs для экситонов, связанных на мелких примесях, ширина линии составляет около 0,1 мэВ [7]. Часто одновременно в одном и том же веществе существуют как свободные, так и связанные экситоны; их можно различить по энергии и ширине линии (например, при 1,4 К в высококачественном GaAs ширина линии $\Delta h\nu = 0,1$ мэВ для связанного экситона против 1 мэВ для свободного экситона).

Фиг. 6.7—6.11 демонстрируют характеристики экситонного излучения. Фиг. 6.7 представляет спектр излучения достаточно чистого InP [12]. Полагают, что пик излучения, обозначенный цифрой 1, обусловлен рекомбинацией свободного экситона, в то время как линии 2, 3, 4 и 5 относятся к рекомбинации связанного экситона, с испусканием соответственно 0, 1, 2 и 3 фононов. Расстояние между линиями 2, 3, 4, 5 составляет 43 мэВ, что соответствует известному значению энергии $1,0$ -фонона. Полосы I и II обусловлены другими процессами, неэкситонного характера.

Фиг. 6.8 характеризует излучательную рекомбинацию связанного экситона в CdS. Пик I_1 — бесфононное излучение. Более



Фиг. 6.7. Зависимость фотолуминесценции InP при 6 К от энергии фотонов [12].

низкоэнергетические пики связаны с испусканием различных фононов в соответствии с обозначениями.

БЕЗЫЗЛУЧАТЕЛЬНАЯ РЕКОМБИНАЦИЯ

Электронно-дырочная пара может рекомбинировать безызлучательно. Безызлучательная рекомбинация является доминирующим процессом во многих полупроводниках. Так, например, вероятность излучательных переходов, рассчитанная для чистого германия из соотношения ван Русбрека — Шокли (см. § 1 гл. 6), соответствует излучательному времени жизни порядка одной секунды, тогда как измеренное время жизни самое большее порядка миллисекунды, а часто меньше микросекунды. Итак, процесс безызлучательной рекомбинации в германии по крайней мере в тысячу раз более вероятен, чем излучательные переходы.

Безызлучательная рекомбинация представляет некоторые трудности для понимания. Термин «излучательная рекомбинация» совершенно ясен — в процессе перехода излучается фотон, в то время как «безызлучательная рекомбинация» — это нечто неопределенное, неявно указывающее на какой-то процесс, при котором не излучается фотон, и оставляющее нашей изобретательности и воображению конструирование подходящей модели. Экспериментальное исследование безызлучательных процессов чрезвычайно трудно, поскольку этот механизм проявляет себя только по отсутствию некоторого другого ожидаемого результата. При этом измеряемыми параметрами являются: эффективность излучения, время жизни носителей и кинетические характеристики рекомбинационного процесса в зависимости от растущей температуры или концентрации носителей. В конечном счете, вывод делается скорее из косвенных данных, чем на основе прямых доказательств.

В этой главе мы рассмотрим несколько рекомбинационных процессов, которые не приводят к излучению фотонов: оже-эффект, поверхностную рекомбинацию и эмиссию фононов. Хотя иногда кажется, что оже-эффект может очень хорошо описать наблюдаемые процессы — это не единственно возможное объяснение. Поэтому область безызлучательной рекомбинации — пока еще открытое поле для дальнейших исследований.