

К. В. ШАЛИМОВА

ФИЗИКА ПОЛУПРОВОДНИКОВ

УЧЕБНИК

*Издание четвертое,
стереотипное*



САНКТ-ПЕТЕРБУРГ • МОСКВА • КРАСНОДАР
2010

ББК 22.379я73

Ш 18

Шалимова К. В.

Ш 18 Физика полупроводников: Учебник. 4-е изд., стер. — СПб.: Издательство «Лань», 2010. — 400 с.: ил. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-0922-8

В учебнике рассмотрены модельные представления о механизме электропроводности, даны основы зонной теории полупроводников и теории колебаний кристаллической решетки, изложена статистика электронов и дырок, рассмотрены механизмы рассеяния носителей заряда, генерация и рекомбинация носителей заряда, диффузия и дрейф неравновесных носителей заряда, изложены контактные и поверхностные явления в полупроводниках, их оптические и фотоэлектрические свойства.

Учебник предназначен для студентов физических и технических специальностей, может быть полезен инженерно-техническим работникам.

ББК 22.379я73

Обложка

А. Ю. ЛАПШИН

*Охраняется законом РФ об авторском праве.
Воспроизведение всей книги или любой ее части
запрещается без письменного разрешения издателя.*

*Любые попытки нарушения закона
будут преследоваться в судебном порядке.*

© Издательство «Лань», 2010
© К. В. Шалимова, наследники, 2010
© Издательство «Лань»,
художественное оформление, 2010

ПРЕДИСЛОВИЕ

В связи со все увеличивающимися темпами развития электроники и повышенным вниманием к фундаментальным и техническим разработкам в области физики полупроводников возникла устойчивая потребность в подготовке специалистов данного направления. Современная литература посвящена преимущественно новейшим достижениям и разработкам, таким как, например, наноэлектроника. При создании актуальных, современных учебных пособий авторам неминуемо приходится меньше внимания уделять методической составляющей издания, грамотному и понятному изложению основ. В результате учебное пособие дает студенту знания и интерес к данному направлению, но не обеспечивает понимания простейших процессов. Возникает необходимость в переиздании методически выверенной литературы, получившей признание многих поколений студентов.

Чтобы восполнить эту потребность, читателям предлагается четвертое издание учебника К. В. Шалимовой «Физика полупроводников».

Книга создавалась в 1960-е годы в период интенсивного развития полупроводниковой электроники в нашей стране. В 1971 г. вышло ее первое издание. Оно не было первым среди книг по физике полупроводников, но выделялось достаточно простым и в то же время строгим изложением материала, охватывающим все разделы, необходимые для студентов, обучающихся по специальностям, связанным с технологией получения полупроводниковых материалов, полупроводниковыми приборами и полупроводниковой электроникой. Уже второе ее издание в 1976 г. получило гриф Министерства высшего и среднего образования СССР и стало использоваться как учебник. Впоследствии книга была еще раз переиздана в 1985 г. Свидетельством ее успешности был перевод на испанский, польский и чешский языки.

За прошедшие с момента последнего издания учебника К. В. Шалимовой 15 лет физика полупроводников обогатилась как исследованиями новых явлений в физике полупроводников и твердого тела, так и исследованиями новых материалов. В первую очередь, вероятно, следует говорить о квантово-размерных эффектах, создании структур квантовых ям, квантовых проволок, квантовых точек и, наконец, сверхрешеток, которые сыграли большую роль в создании новых приборов и совершенствовании параметров уже известных. Более того, переход на наноуровень в размерах исследуемых структур поставил ряд новых вопросов в понимании происходящих в них физических процессов.

Все эти аспекты выходят за рамки данной книги. Несмотря на это, учебник К. В. Шалимовой остается одним из самых востребованных в этой области. Мы надеемся, что новое издание найдет большое число читателей среди преподавателей и студентов и по-своему обогатит образовательный процесс.

СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

- a — постоянная решетки, радиус Бора, ускорение
 a, b, c — полуоси эллипсоида
 $B, \mathcal{B}$ — магнитная индукция
 C — емкость, теплопроводность
 c — скорость света в вакууме
 D — биполярный коэффициент диффузии; прозрачность потенциального барьера
 D_n, D_p — коэффициент диффузии электронов, дырок
 $\mathcal{E}, \mathcal{E}$ — напряженность электрического поля
 $\mathcal{E}_{кр}$ — критическая напряженность электрического поля
 $\mathcal{E}_M$ — максимальная напряженность электрического поля
 E — энергия
 E_c, E_v — энергии, соответствующие дну зоны проводимости и потолку валентной зоны
 $E_a, E_d, E_{акс}$ — энергии ионизации акцептора, донора, экситона
 E_g — ширина запрещенной зоны
 $E_{и}$ — энергия ионизации
 E_p — энергия фонона
 E_T — порог внешнего фотоэффекта
 E_t — энергия уровня захвата
 e — заряд электрона; деформация
 $\hat{H}$ — оператор Гамильтона
 H, H — напряженность магнитного поля
 h — постоянная Планка; $\hbar = h/2\pi$
 F, F — сила
 F — уровень Ферми (энергия Ферми)
 F_n, F_p — квазиуровень Ферми для электронов и дырок
 f — функция Ферми—Дирака; неравновесная функция распределения
 f_0, f_{op} — равновесная функция распределения для электронов и дырок
 G — скорость генерации
 G_s — поверхностная проводимость
 g — фактор спинового вырождения; вероятность поглощения
 I — ток, интенсивность света
 i, j, k — единичные векторы
 J, J — плотность тока
 J_s — плотность тока насыщения
 k, k — волновой вектор носителя заряда; модуль волнового вектора
 k — постоянная Больцмана
 k_n, k_p — вероятность захвата электрона, дырки
 $\mathcal{L}$ — длина; толщина слоя объемного заряда
 L — длина; диффузионная длина; длина затягивания
 L_n, L_p — длина диффузии электронов и дырок
 l — длина свободного пробега; длина
 l_3 — длина экранирования
 M — масса ядра; масса атома (иона)
 m_0 — масса свободного электрона
 m^* — эффективная масса носителя заряда
 m_n^*, m_p^* — эффективная масса электрона, дырки

m_{dn}^*, m_{dp}^* — эффективная масса плотности состояний электронов, дырок
 N — плотность квантовых состояний; концентрация
 N_a, N_d — концентрация акцепторов и доноров
 N_c, N_v — эффективная плотность состояний в зоне проводимости, валентной зоне
 N_t — концентрация ловушек захвата
 N_p — плотность фононов
 N_s — плотность поверхностных состояний
 n — неравновесная концентрация электронов; квантовое число
 n_a, n_d — концентрация электронов на акцепторной, донорной примеси
 n_0 — равновесная концентрация электронов
 n_i — собственная концентрация носителей заряда
 n — комплексный показатель преломления
 $\bar{n}$ — показатель преломления среды
 P, P — импульс электрона; поляризуемость среды; вероятность перехода
 p, p — квазиимпульс носителя заряда
 p — неравновесная концентрация дырок; показатель степени
 p_a, p_d — концентрация ионизованной акцепторной, донорной примеси
 p_0 — равновесная концентрация дырок
 Π — коэффициент Пельтье
 Q, Q — квазиимпульс решетки; объемный заряд; полный заряд; количество тепла
 Q_{ss} — поверхностный заряд
 q, q — волновой вектор фонона; модуль волнового вектора фонона
 R, R, r, r — радиус-вектор
 R — сопротивление; коэффициент Холла, скорость рекомбинации; коэффициент отражения
 Q_{sp}^0, R_{sp}^0 — объемный заряд в приповерхностной области полупроводника
 R_n^0, R_p^0 — плотность равновесного излучения, поглощения
 r, r — радиус-вектор электрона
 s — площадь; скорость поверхностной рекомбинации
 T — температура; время; кинетическая энергия; период; пропускание
 t — время
 U — потенциальная энергия; разность потенциалов
 u — смещение
 u_v — плотность излучения
 V — потенциал кристаллической решетки; объем кристалла
 V_B — объем зоны Бриллюэна
 v, v — дрейфовая скорость
 v_0 — средняя скорость теплового движения
 v_D — диффузионная скорость
 v_d — скорость домена
 v_{zv} — скорость звука
 v_{gr} — групповая скорость
 $v_{ф}$ — фазовая скорость
 W — поток тепловой энергии
 W — вероятность
 w — энергия активации
 Y — безразмерный поверхностный потенциал
 α — термический коэффициент сопротивления; коэффициент ударной ионизации; удельная термо-ЭДС; коэффициент поглощения
 β — температурный коэффициент; коэффициент квазиупругой силы; квантовый выход
 γ — коэффициент рекомбинации; коэффициент пропорциональности
 $\Delta n, \Delta p$ — концентрация избыточных электронов, дырок
 δ — относительное изменение энергии; дельта-функция; толщина туннельного барьера
 ϵ_r, ϵ_0 — относительная диэлектрическая проницаемость среды; электрическая постоянная
 η — приведенный уровень Ферми; импульс фотона

Θ — температура Дебая; угол
 ϕ — угол
 κ — теплопроводность
 κ_e, κ_L — электронная, решеточная теплопроводность
 λ — длина волны
 μ, μ_d — дрейфовая подвижность носителей заряда
 μ_D — амбиполярная диффузионная подвижность
 μ_n, μ_p — подвижность электронов, дырок
 $\mu_{\mathcal{E}}$ — амбиполярная дрейфовая подвижность
 μ_n — холловская подвижность носителей заряда
 μ_r, μ_0 — относительная магнитная проницаемость среды; магнитная постоянная
 ν — частота; постоянная спада
 ξ — уровень Ферми, отсчитанный от E_c ; химический потенциал
 ρ — удельное сопротивление; линейная плотность; объемный заряд
 σ — удельная проводимость; эффективное сечение
 τ — среднее время свободного пробега; время релаксации; постоянная времени; время жизни
 τ_n, τ_p — время жизни электронов, дырок
 τ_μ — время релаксации Максвелла
 τ_T — коэффициент Томсона
 φ — функция; угол; угол Холла; потенциал; контактная разность потенциалов
 Φ — функция; работа выхода
 χ — энергия сродства к электрону; работа выхода; показатель поглощения
 Ψ — функция; поверхностный потенциал
 ω — угловая частота
 ω_c — циклотронная частота

ГЛАВА ПЕРВАЯ

ПОЛУПРОВОДНИКИ.

ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ЭЛЕКТРОПРОВОДНОСТИ

1.1. КЛАССИФИКАЦИЯ ВЕЩЕСТВ ПО УДЕЛЬНОЙ ЭЛЕКТРИЧЕСКОЙ ПРОВОДИМОСТИ. ПОЛУПРОВОДНИКИ

Все вещества по электрофизическим свойствам могут быть разделены на три больших класса: металлы, полупроводники и диэлектрики. Наиболее просто, казалось бы, классифицировать вещества по удельному электрическому сопротивлению ρ . У металлов оно находится в пределах $10^{-6} \div 10^{-4}$ Ом·см (например, для серебра при комнатной температуре $\rho = 1,58 \cdot 10^{-6}$ Ом·см, для сплава нихром $\rho = 1,05 \cdot 10^{-4}$ Ом·см). Вещества с удельным сопротивлением от 10^{-4} до 10^{10} Ом·см были отнесены к полупроводникам (например, удельное сопротивление сернистого кадмия при комнатной температуре в зависимости от технологии его изготовления лежит в пределах от 10^{-3} до 10^{12} Ом·см, германия от 10^{-4} до 47 Ом·см, а кремния от 10^{-4} до $2,5 \cdot 10^5$ Ом·см). Наконец, вещества с удельным сопротивлением более 10^{10} Ом·см являются диэлектриками (например, при 200 °С для слюды в зависимости от ее состава $\rho = 10^{13} \div 10^{16}$ Ом·см, для стекла $\rho = 10^8 \div 10^{15}$ Ом·см).

Из приведенных примеров следует, что при переходе от одного класса вещества к другому значения удельного сопротивления перекрываются. Поэтому значения удельного сопротивления не могут служить в качестве однозначного критерия для классификации веществ. Как следует из рис. 1.1 и 1.2, металлы и полупроводники имеют различного вида температурные зависимости удельного сопротивления и удельной электрической проводимости. У металлов с ростом температуры ρ увеличивается пропорционально абсолютной температуре T , т. е.

$$\rho = \rho_0 (1 + \alpha t) = \frac{\rho_0}{T_0} T, \quad (1.1)$$

где ρ_0 — удельное сопротивление данного металла при 0 °С; α — температурный коэффициент сопротивления, равный $1/273$; $T_0 = 273$ К.

Для металлов, как следует из рис. 1.2, *а*, характерным является наличие отрицательного температурного коэффициента удельной электрической проводимости.

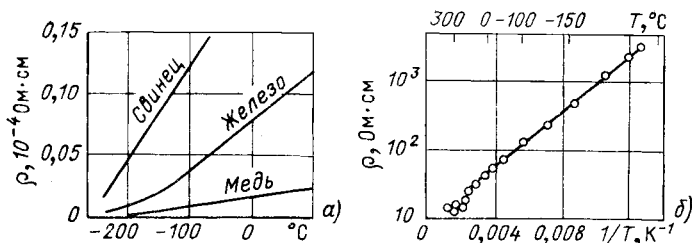


Рис. 1.1. Зависимость удельного сопротивления металлов (а) и кремния (б) от температуры

У полупроводников характер температурной зависимости удельного сопротивления и проводимости иной, для некоторого интервала температур они определяются зависимостями вида:

$$\rho = \rho_0 e^{\beta/T}; \quad (1.2)$$

$$\sigma = \sigma_0 e^{-\beta/T}, \quad (1.3)$$

где ρ_0 , σ_0 , β — некоторые постоянные для данного интервала температур, характерные для каждого полупроводникового вещества.

Такие зависимости удельного сопротивления и удельной проводимости от температуры имеют так называемые не вырожденные полупроводники. У них согласно рис. 1.2, б температурный коэффициент удельной проводимости положительный, т. е.

$$\frac{\Delta\sigma}{\Delta T} = \frac{\sigma_2 - \sigma_1}{T_2 - T_1} > 0.$$

Казалось бы, что теперь вопрос о различии полупроводников и металлов решен знаком температурного коэффициента удельной проводимости. Однако выбор его в качестве определяющего критерия осложнен тем, что в некотором интервале температур полупроводник может вести себя подобно металлу. Поэтому по знаку температурного коэффициента

удельной проводимости не всегда можно установить принадлежность вещества к классу полупроводников.

Ответить на этот вопрос можно, если проследить, как изменяется проводимость вещества при понижении температуры. Как следует из рис. 1.2, а,

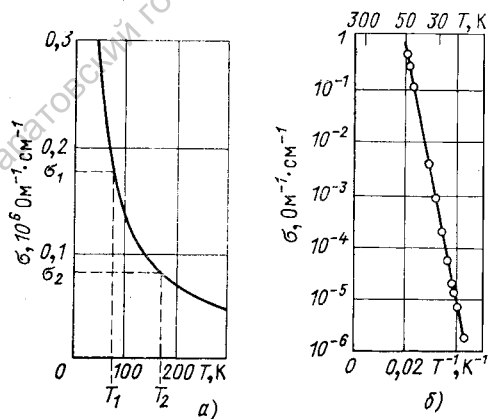


Рис. 1.2. Зависимость удельной проводимости свинца (а) и кремния (б) от температуры

с понижением температуры удельная проводимость металлов растет, что возможно лишь потому, что в металле независимо от температуры всегда имеются свободные электроны. У полупроводников, наоборот, удельная проводимость уменьшается при понижении температуры (рис. 1.2, б). Из этого следует, что в полупроводнике свободные носители заряда возникают в результате теплового возбуждения. Эти носители заряда называют **тепловыми** или **равновесными**. Опыт показывает, что появление свободных носителей заряда в полупроводнике имеет место также при освещении, облучении ядерными частицами. Возникающие в этих случаях носители заряда принято называть **неравновесными**. Процесс образования как равновесных, так и неравновесных носителей заряда очень сильно зависит от структуры полупроводникового материала и наличия в нем примеси.

Следовательно, *полупроводники — это такие материалы, которые при комнатной температуре имеют удельную проводимость в интервале от 10^{-10} до 10^4 Ом $^{-1}$ ·см $^{-1}$, зависящую в сильной степени от структуры вещества, вида и количества примеси и от внешних условий: температуры, освещения, облучения ядерными частицами электрического и магнитного полей.*

Существует, как известно, два типа проводников электрического тока: электронные и ионные. Металлы — электронные проводники. У ионных проводников ток переносится ионами вещества, вследствие чего состав ионного проводника меняется при прохождении через него электрического тока. Типичные представители ионных проводников — электролиты.

Полупроводниковые вещества также могут быть как электронными, так и ионными. В данной книге будут рассмотрены физические свойства только электронных полупроводников.

К электронным полупроводникам относится огромное количество самых различных веществ. Полупроводниками являются как простые вещества: бор В, кремний Si, фосфор Р, сера α -S, германий Ge, мышьяк As, серое олово α -Sn, сурьма β -Sb, селен Se, теллур Te, так и многие сложные химические соединения. В качестве примера в табл. 1.1 приведены некоторые соединения элементов периодической системы Менделеева (рис. 1.3), обладающие полупроводниковыми свойствами. В таблице указаны кристаллическая структура, постоянная решетки и другие характеристики этих важнейших полупроводников.

Кроме неорганических веществ полупроводниковыми свойствами обладают также и некоторые органические вещества, такие как фталоцианины и полициклические ароматические углеводороды (например, бензол, нафталин, антрацен, нафтацен, коронен и др.).

Особенности электрофизических свойств полупроводников обусловили их широкое применение для создания самых различных приборов. Полупроводниковые приборы используются для выпрямления переменного тока (диоды), преобразования или смещения частот, усиления и генерации электрических колебаний (транзисторы, СВЧ-диоды, диоды Ганна, лавинно-пролетные диоды, тун-

Т а б л и ц а 1.1. Кристаллическая структура важнейших полупроводников

Типы полупроводников	Элемент или соединение	Название	Кристаллическая структура	Постоянная решетки при 300 К, нм	Ширина запрещенной зоны при 0 К, эВ	Относительная диэлектрическая проницаемость
Элементарные полупроводники	C	Алмаз	Алмаз	0,356679	5,51	5,5
	Ge	Германий		0,565748	0,74	16
	Si	Кремний		0,543086	1,16	11,8
	Sn	Серое олово		0,64892	0,08	
$A^{IV}-B^{IV}$	SiC	Карбид кремния	Цинковой обманки (кубическая)	0,4358	2,68	10
$A^{III}-B^V$	AlSb	Антимонид алюминия	Цинковой обманки	0,61355	1,60	11
	GaSb	Антимонид галлия		0,60955	0,81	15
	GaAs	Арсенид галлия		0,56534	1,52	12
	GaP	Фосфид галлия		0,54505	2,40	10
	InSb	Антимонид индия		0,64787	0,235	18
	InAs	Арсенид индия		0,60585	0,43	12,5
	InP	Фосфид индия		0,58687	1,42	12,1
$A^{II}-B^{VI}$	CdS	Сульфид кадмия	Цинковой обманки Вюрцита (гексагональная) Цинковой обманки Вюрцита Цинковой обманки Вюрцита	0,5832	2,487	10
	CdS	Сульфид кадмия		а0,4136 с0,6713	2,58	
	CdSe	Селенид кадмия		0,605	1,85	10
	ZnO	Оксид цинка		а0,32496 с0,52065	3,43	8,5
	ZnS	Сульфид цинка		0,5409	3,83	8,3
	ZnS	Сульфид цинка		а0,3826 с0,626	3,91	8,7
$A^{IV}-B^{VI}$	PbS PbTe	Сульфид свинца Теллурид свинца	Кубическая »	0,5935 0,6460	0,29 0,165	170 250

ПЕРИОДИЧЕСКАЯ СИСТЕМА ЭЛЕМЕНТОВ Д. И. МЕНДЕЛЕЕВА															
I		II		III		IV		V		VI		VII		VIII	
H		Be		B		C		N		O		(H)		He	
1	100794	4	90018	5	1081	6	1201	7	140067	8	15393,	9	18398403	10	400760
ВОДОРОД		БЕРИЛЛИЙ		БОР		УГЛЕРОД		АЗОТ		КИСЛОРОД		ФТОР		ГЕЛИЙ	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li		Be		B		C		N		O		(H)	
1		2		3		4		5		6		7		8	
H		Li													

нельные диоды), преобразования тепловой энергии и энергии излучения в электрическую (термоэлементы и фотоэлементы), для преобразования электрической энергии в световую (фосфоры, светоизлучающие диоды, лазеры). Полупроводниковые приборы применяются также в качестве «датчиков» для измерения температуры (термисторы), регистрации светового и корпускулярного излучения (фоторезисторы и дозиметры), для измерения давления (тензодатчики) и магнитных полей (датчики Холла).

В основе работы каждого прибора лежат определенные физические процессы и явления. Другими словами, физика полупроводников представляет собой теоретическую основу полупроводниковой электроники. Разработка полупроводниковых материалов, создание на их основе новых видов приборов, научные исследования в этой области, постановка высококачественного производства полупроводниковых приборов и интегральных схем, разработка новых устройств для разнообразных условий применения — все это возможно при условии серьезных знаний в области физики полупроводников.

1.2. МОДЕЛЬНЫЕ ПРЕДСТАВЛЕНИЯ О МЕХАНИЗМЕ ЭЛЕКТРОПРОВОДНОСТИ СОБСТВЕННЫХ ПОЛУПРОВОДНИКОВ

Рассмотрим механизм электропроводности одноатомных полупроводников, например кремния и германия. Кремний и германий находятся в одной подгруппе периодической системы Менделеева с углеродом. Электроны распределены у них по состояниям следующим образом:

$$\text{C}^{(6)} (1s^2 2s^2 2p^2);$$

$$\text{Si}^{(14)} (1s^2 2s^2 2p^6 3s^2 3p^2);$$

$$\text{Ge}^{(32)} (1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^2).$$

Внешняя электронная оболочка у этих атомов заполнена частично, она содержит четыре электрона: 2 в s - и 2 в p -состояниях.

При образовании кристалла, например кремния, четыре валентных электрона каждого атома из состояния $3s^2 3p^2$ переходят в гибридное sp^3 -состояние с антипараллельными спинами и образуют четыре ковалентные связи. В результате каждый атом окружен четырьмя ближайшими соседями и находится в центре тетраэдра. Такая кристаллическая решетка является кубической и получила название решетки типа алмаза (рис. 1.4). В ней четыре внешних электрона каждого атома участвуют в образовании ковалентных связей (по два электрона в каждой). Эти связи на рис. 1.4 представлены в виде двух направляющих, соединяющих два ближайших атома. На рис. 1.5, а дано двумерное изображение ковалентных связей в решетке кремния. Здесь в узле решетки находится ион кремния с зарядом $+4$, которому принадлежат четыре валентных электрона. Валентные электроны, обеспечивающие ковалентную связь, на рис. 1.5, а представлены в виде черных точек.

В полупроводнике с идеальной структурой, изображенной на рис. 1.5, а, все электроны находятся в связанном состоянии. Если такой полупроводник поместить во внешнее электрическое поле, то электрический ток не может возникнуть, так как все ковалентные связи в решетке завершены и свободных носителей заряда нет.

Допустим, что при воздействии каких-либо возмущений, например теплоты, в полупроводнике произошел разрыв ковалентной

связи и электрон стал свободным. Процесс превращения связанного электрона в свободный электрон носит название генерации. При уходе электрона ковалентная связь будет незавершенной. Незавершенная связь будет иметь избыточный положительный заряд, поскольку он теперь некомпенсирован зарядом электрона. *Вакантное место в ковалентной связи получило название дырки.* В целом кристалл остается электронейтральным, так как каждому образовавшемуся положительному заряду в связи — дырке соответствует свободный электрон. На рис. 1.5, б свободные электроны и дырки изображены соответственно черными и светлыми кружочками. Свободный электрон может занять вакантное место в ковалентной связи и перейти в связанное состояние. Процесс превращения сво-

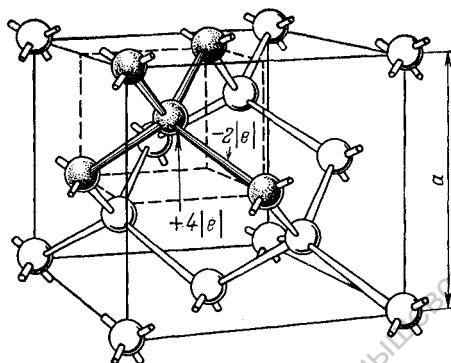


Рис. 1.4. Кристаллическая решетка типа алмаза:

a — постоянная решетки

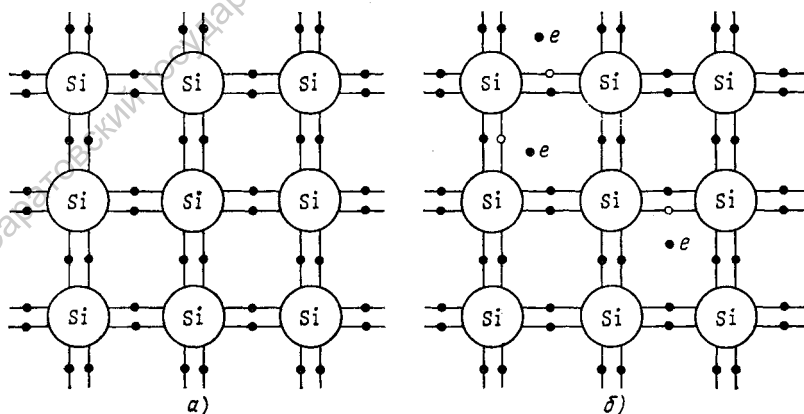


Рис. 1.5. Двумерное представление расположения связей в решетке кремния (собственный полупроводник)

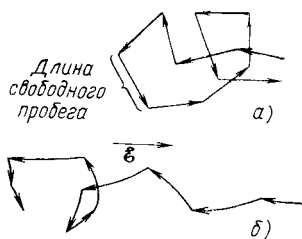


Рис. 1.6. Схема движения свободного электрона при отсутствии (а) и наличии внешнего электрического поля (б)

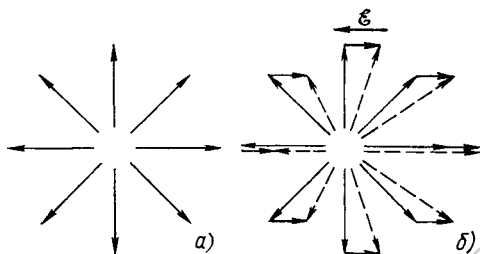


Рис. 1.7. Схематическое изображение скоростей электронов проводимости при отсутствии (а) и наличии внешнего электрического поля (б)

бодного электрона в связанный электрон носит название рекомбинации.

Следовательно, *собственный полупроводник* — идеально чистое вещество, в котором в результате разрыва ковалентных связей образуется равное количество свободных электронов и дырок. При комнатной температуре концентрация свободных электронов и дырок в германии составляет примерно 10^{13} см^{-3} , а у кремния — около 10^{10} см^{-3} . Увеличение температуры полупроводника приводит к возрастанию количества разорванных ковалентных связей, а следовательно, к росту концентрации свободных электронов и дырок в полупроводнике.

В отсутствие внешнего электрического поля свободный электрон, совершая тепловое движение в кристалле полупроводника, будет сталкиваться с дефектами кристаллической решетки, в результате чего меняется направление его движения. В силу этого тепловое движение свободного электрона является беспорядочным, как это изображено на рис. 1.6, а.

Вакантное место в ковалентной связи — дырка может быть заполнена электроном, перешедшим вследствие теплового возбуждения с соседней насыщенной ковалентной связи. При таком переходе от атома к атому дырка также будет совершать хаотическое движение.

Таким образом, тепловое движение свободных носителей заряда в кристалле полупроводника является беспорядочным. При этом каждый из носителей заряда описывает сложную траекторию движения (рис. 1.6, а). Расстояние, проходимое свободным носителем заряда между двумя столкновениями, называют *длиной свободного пробега*, а *усредненное значение всех отрезков пути, пройденного носителем, есть средняя длина свободного пробега*. Соответственно время между двумя соударениями и его усредненное значение называют *временем свободного пробега* и *средним временем свободного пробега*.

Средняя длина свободного пробега l и среднее время свободного пробега τ связаны соотношением

$$l = v_0 \tau, \quad (1.4)$$

где v_0 — средняя скорость теплового движения свободного носителя заряда (среднее значение модуля скорости).

В полупроводниках при комнатной температуре скорость теплового движения электронов порядка 10^7 см/с.

Поскольку тепловое движение свободных носителей заряда — электронов и дырок — является беспорядочным, то их средняя скорость, рассматриваемая как векторная величина, равна нулю. Это означает, что каждому движущемуся свободному электрону (дырке) можно сопоставить свободный электрон (дырку), перемещающийся в противоположную сторону с такой же скоростью (рис. 1.7, а). Следовательно, вследствие беспорядочного теплового движения количество свободных электронов и дырок, движущихся в любом направлении, в среднем равно числу электронов и дырок, движущихся в противоположном направлении. Поэтому в отсутствие внешнего электрического поля суммарный заряд, переносимый свободными электронами и дырками в любом направлении, равен нулю, а поэтому беспорядочное тепловое движение носителей заряда не дает тока.

Если собственный полупроводник находится во внешнем электрическом поле, то под воздействием этого поля свободные электроны приобретут скорость, направленную против поля. Благодаря этому у электронов, тепловое движение которых происходило против направления внешнего электрического поля, скорость увеличится, а у электронов, движущихся в направлении поля, уменьшится.

В результате вся совокупность свободных электронов получает некоторую скорость движения в направлении, противоположном направлению внешнего электрического поля (рис. 1.7, б). Изменение скорости движения электронов скажется на их кинетической энергии. При столкновении с дефектами кристаллической решетки электроны передают решетке всю энергию, приобретенную в поле, вследствие чего они придут в тепловое равновесие с решеткой. После столкновения электроны, совершая хаотическое тепловое движение, опять будут иметь направленное движение во внешнем электрическом поле.

Таким образом, фактическое движение электрона в кристалле складывается из беспорядочного теплового и упорядоченного движения, вызванного действием внешнего электрического поля (рис. 1.6, б). В результате этого происходит медленное перемещение всей совокупности свободных электронов с некоторой средней скоростью. Направленное движение совокупности свободных носителей заряда во внешнем электрическом поле носит название дрейфа, а скорость их направленного движения называют дрейфовой. На рис. 1.7, б горизонтальными линиями изображены векторы дрейфовой скорости, а пунктиром — векторы суммарной скорости теплового движения и дрейфа. В этом случае средняя скорость движения свободных электронов уже не равна нулю и через полупроводник в направлении электрического поля потечет ток, обусловленный свободными электронами.

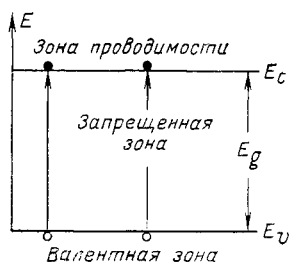


Рис. 1.8. Схематическое изображение энергетических зон собственного полупроводника

Электроны насыщенных связей при переходе в вакантное место в связи под действием внешнего электрического поля будут перемещаться против направления поля. Тем самым вакантное место в ковалентной связи — дырка будет перемещаться, но по направлению внешнего электрического поля, что равносильно перемещению по полю положительного заряда. Механизм электропроводности, обусловленный движением электронов по свободным местам в ковалентных связях, получил название дырочной электропроводности.

Таким образом, в чистом полупроводнике, не содержащем примесей, осуществляется электронная и дырочная электропроводность. Следовательно, электрический ток в собственном полупроводнике определяется двумя составляющими — электронным и дырочным токами, текущими в одном направлении.

Электропроводность собственного полупроводника можно объяснить, если исходить из энергетических представлений. Электроны в изолированном атоме, как показывает квантовая теория, обладают дискретными значениями энергии. При этом на каждом энергетическом уровне согласно принципу Паули может находиться не более двух электронов, которые должны иметь противоположно направленные спины. При образовании полупроводникового вещества, т. е. при сближении атомов на расстояние примерно 10^{-8} см, будет происходить перекрытие волновых функций атомарных электронов.

Благодаря этому энергетический уровень валентных электронов превращается в зону. Эта зона носит название валентной. Из вышележащего уровня возбуждения атомов образуется зона, которую называют зоной проводимости. Каждая из этих зон занимает определенную область энергии. Зоны дозволённых энергий отделены друг от друга интервалом, называемым запрещённой зоной, или энергетической щелью. На рис. 1.8, на котором изображена энергетическая диаграмма собственного полупроводника, E_c соответствует минимальной энергии, которую могут иметь свободные электроны в кристалле, величина E_v — максимальная энергия электронов валентной зоны, E_g — ширина запрещённой зоны, значение которой определяется природой химической связи, а также зависит от атомов, образующих кристаллическую решетку¹.

Как известно, плотность тока есть заряд, переносимый электронами за единицу времени через единичную площадку, перпендику-

Индексы *c*, *v* и *g* происходят от английских слов: conduction — проводимость; valence — валентный; gap — промежуток (щель).

лярную их скорости v . Поэтому плотность тока J , обусловленного движением одного электрона в объеме Ω ,

$$J = - \frac{e}{\Omega} v. \quad (1.5)$$

Предположим, что все возможные квантовые состояния в энергетической зоне заняты электронами. В этом случае плотность электрического тока для всей системы электронов

$$J = - \frac{e}{\Omega} \sum_{\substack{s \\ \text{По всей} \\ \text{зоне}}} v_s = 0, \quad (1.6)$$

так как каждому электрону со скоростью v можно сопоставить электрон, со скоростью $-v$, порождающий равный по величине, но противоположный по направлению ток.

В собственном полупроводнике при разрыве ковалентной связи, например, за счет теплоты появляются свободный электрон и вакантное место в валентной связи — дырка. В схеме энергетических зон, изображенной на рис. 1.8, это равнозначно переходу электрона из валентной зоны в зону проводимости. В этом случае все квантовые состояния валентной зоны будут заняты электронами, за исключением одного состояния. Скорость движения носителя заряда при переходе в свободное состояние в ковалентной связи обозначим через v_i . Тогда суммарная плотность тока всех электронов валентной зоны запишется в виде

$$J = - \frac{e}{\Omega} \sum_{s \neq i} v_s = - \frac{e}{\Omega} \sum_{\substack{s \\ \text{По всей} \\ \text{зоне}}} v_s + \frac{e}{\Omega} v_i. \quad (1.7)$$

Но сумма плотностей токов в правой части (1.7) по всем состояниям в зоне равна нулю. Поэтому

$$J = \frac{e}{\Omega} v_i. \quad (1.8)$$

Таким образом, суммарный ток всех электронов в валентной зоне эквивалентен току одного электрона, если поместить последний в вакантное место в ковалентной связи и приписать ему положительный заряд $+e$. Такое вакантное состояние ковалентной связи принято называть положительной дыркой, или просто дыркой.

При обычных температурах в зоне проводимости собственного полупроводника всегда имеется некоторое количество электронов, заброшенных туда из валентной зоны путем теплового возбуждения, а в его валентной зоне будет равное количество положительных дырок. При наличии внешнего электрического поля свободные электроны и дырки будут осуществлять перенос электрического заряда.

1.3. МОДЕЛЬНЫЕ ПРЕДСТАВЛЕНИЯ О МЕХАНИЗМЕ ЭЛЕКТРОПРОВОДНОСТИ ПРИМЕСНЫХ ПОЛУПРОВОДНИКОВ

Рассмотрим теперь механизм электропроводности полупроводника с решеткой типа алмаза, в котором один из атомов замещен атомом элемента V группы, например мышьяка в решетке кремния. Полупроводник, имеющий примеси, называют *примесным*, а его электропроводность, обусловленную наличием в кристалле примеси, — *примесной*. У атома мышьяка пять валентных электронов расположены в $4s$ - и $4p$ -состояниях: $As^{(33)} (1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^3)$. В решетке кремния четыре валентных электрона атома мышьяка вместе с четырьмя электронами ближайших атомов кремния участвуют в образовании ковалентной связи, как это схематически представлено на рис. 1.9, а. Пятый электрон мышьяка, не принимающий участие в образовании ковалентной связи, слабо связан с атомом мышьяка, так как он находится в электрическом поле электронов окружающих атомов кремния. При низких температурах пятый электрон локализован около атома мышьяка, но при повышенных температурах он может быть отщеплен от примеси и будет свободно перемещаться по кристаллу. В этом случае также соблюдается электронейтральность кристалла, так как атом мышьяка, отдавший пятый электрон, будет теперь положительным ионом.

Наряду с ионизацией примеси может происходить и ионизация атомов основного вещества. Но в области температур ниже той, при которой имеет место значительная собственная электропроводность, количество электронов, оторванных от примеси, будет значительно больше количества электронов и дырок, образовавшихся в результате разрыва ковалентных связей. В силу этого *доминирующую роль в проводимости полупроводника будут играть электроны, поэтому их будем называть основными носителями заряда*, а

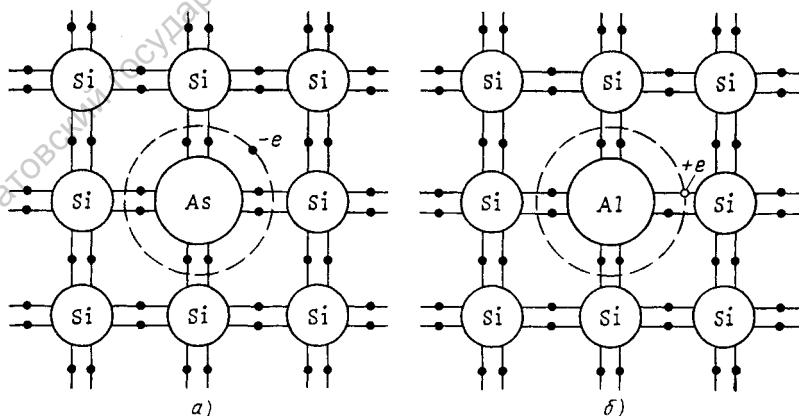


Рис. 1.9. Двумерное представление расположения связей в решетке донорного (а) и акцепторного (б) полупроводников

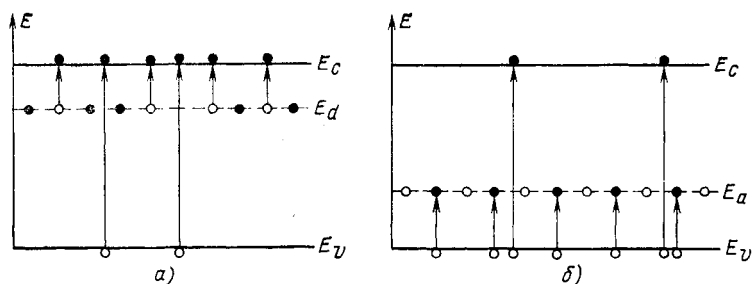


Рис. 1.10. Энергетическая диаграмма донорного (а) и акцепторного (б) полупроводников

дырки — неосновными носителями заряда. Такой полупроводник называют электронным или *n*-типа, а примесь, дающую электроны, донорной или примесью *n*-типа.

Наличие примеси в кристалле полупроводника будет характеризоваться появлением локальных уровней, лежащих в запрещенной зоне. Так как при ионизации атома мышьяка образуется свободный электрон и для его возникновения требуется значительно меньшая энергия, чем для разрыва ковалентной связи кремния, энергетический уровень донорной примеси E_d должен располагаться в запрещенной зоне в непосредственной близости к краю зоны проводимости (рис. 1.10, а).

Допустим, что в качестве примеси в кристаллическую решетку полупроводника с ковалентной связью внесены атомы элемента третьей группы периодической системы Менделеева, например алюминий в решетке кремния. Поскольку высшая валентность алюминия равна трем: $\text{Al}^{(13)} (1s^2 2s^2 2p^6 3s^2 3p)$, то одна связь атома кремния будет не завершена (рис. 1.9, б).

В незаполненную связь около атома алюминия в результате теплового возбуждения может перейти электрон от соседнего атома кремния. При этом образуются отрицательный ион алюминия и свободная дырка, перемещающаяся по связям кремния и, следовательно, принимающая участие в проводимости полупроводника. *Примеси, захватывающие электроны, называют акцепторными.* Для образования свободной дырки за счет перехода электрона от атома кремния к атому акцепторной примеси требуется значительно меньше энергии, чем для разрыва ковалентной связи кремния. В силу этого количество дырок значительно больше количества свободных электронов и поэтому в таком полупроводнике основными носителями заряда будут дырки, а электроны — неосновными носителями заряда. Полупроводник с акцепторной примесью носит название дырочного или *p*-типа.

На энергетической диаграмме, изображенной на рис. 1.10, б, акцепторная примесь имеет в запрещенной зоне энергетический уровень E_a , расположенный на небольшом расстоянии над потолком валентной зоны. При ионизации акцепторной примеси проис-

ходит переход электрона из валентной зоны на уровень E_a , а в валентной зоне появляется вакантная ковалентная связь — дырка, которая и является свободным носителем заряда.

В полупроводниках могут одновременно присутствовать как донорная, так и акцепторная примеси.

1.4. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ЭЛЕКТРОПРОВОДНОСТИ ПОЛУПРОВОДНИКОВ

Проведем подсчет плотности тока для донорного полупроводника, электроны проводимости которого будем рассматривать как идеальные частицы, не имеющие собственного объема и не взаимодействующие друг с другом. Так как по классической теории радиус электрона $r_0 \approx 10^{-13}$ см, то при концентрации их $n \approx 10^{22}$ см $^{-3}$ объем электронов составляет $\frac{4}{3} \pi r_0^3 n \approx 10^{-17}$ объема вещества и первое предположение вполне оправдано. Квантовая теория твердого тела, как это будет показано в гл. 2, в известном приближении позволяет рассматривать электроны в кристалле как невзаимодействующие частицы.

Предположим, что концентрация электронов проводимости n (количество свободных электронов в 1 см 3 полупроводника), а скорость их дрейфового движения v . Поскольку плотность тока есть заряд, проходящий в единицу времени через единичное сечение площадки, перпендикулярной направлению скорости движения электронов, то

$$\mathbf{J} = -en\mathbf{v}. \quad (1.9)$$

Определение скорости дрейфа электронов проведем с учетом следующих предположений. Пусть dt/τ есть вероятность того, что электрон за время dt испытает столкновение (рассеяние). Кроме того, будем считать, что вероятность столкновения в единицу времени $1/\tau$ не зависит от времени, т. е. τ есть некоторая постоянная величина.

Количество столкновений для n частиц за время dt соответственно будет равно $n dt/\tau$. Следовательно, за время dt концентрация носителей заряда, движущихся в данном направлении, уменьшится в результате рассеяния на величину

$$-dn = n \frac{dt}{\tau}. \quad (1.10)$$

Решая уравнение (1.10) относительно n , получаем количество электронов, не испытавших к моменту времени t соударения:

$$n(t) = n_0 e^{-t/\tau}, \quad (1.11)$$

где $n = n_0$ при $t = 0$.

Внешнее электрическое поле напряженности $\mathcal{E}$ сообщит электрону с массой m ускорение

$$a = \frac{e\mathcal{E}}{m}. \quad (1.12)$$

За время свободного пробега t электрон приобретает дрейфовую скорость

$$v_t = at = \frac{e\tau}{m} \mathcal{E} \quad (1.13)$$

и пройдет путь

$$x = \frac{e\mathcal{E}}{2m} t^2. \quad (1.14)$$

Расстояние, которое пройдут все электроны по направлению x , совпадающему с направлением внешнего электрического поля, будет:

$$X = \frac{en_0\tau^2\mathcal{E}}{2m} \int_0^\infty \left(\frac{t}{\tau}\right)^2 e^{-t/\tau} \frac{dt}{\tau} = \frac{en_0\tau^2\mathcal{E}}{m}. \quad (1.15)$$

Если n_0 электронов имеют среднее время свободного пробега $\bar{t}$, то

$$T = n_0 \bar{t}. \quad (1.16)$$

Есть определенная вероятность того, что среди электронов имеются такие, которые обладают одним и тем же временем свободного пробега t . Это электроны, испытавшие соударения в момент времени от t до $t + dt$. Количество таких электронов согласно (1.10) составит $n dt/\tau$, время их движения будет $tn dt/\tau$. Тогда, интегрируя это выражение по всем временам свободного пробега, которые вследствие случайного характера столкновений могут принимать значения от 0 до ∞ , определим время движения электронов

$$T = \int_0^\infty \frac{tn dt}{\tau}. \quad (1.17)$$

Воспользовавшись выражениями (1.16), (1.17) и (1.11), получим:

$$\bar{t} = \frac{1}{n_0} \int_0^\infty \frac{tn dt}{\tau} = \frac{1}{n_0} \int_0^\infty \tau n_0 \frac{t}{\tau} e^{-t/\tau} \frac{dt}{\tau} = \tau. \quad (1.18)$$

Таким образом, τ есть среднее время свободного пробега, т. е. среднее время движения электронов между двумя соударениями.

Следовательно, скорость дрейфа электронов

$$v = \frac{X}{T} = \frac{e\tau\mathcal{E}}{m}, \quad (1.19)$$

т. е. она пропорциональна напряженности электрического поля, времени свободного пробега и обратно пропорциональна массе электрона.

Параметр, связывающий дрейфовую скорость носителей заряда с напряженностью электрического поля, называют подвижностью носителей заряда. Обозначим его μ , тогда

$$v = \mu \mathcal{E}, \quad (1.20)$$

откуда

$$\mu = v/\mathcal{E} = e\tau/m, \quad (1.21)$$

т. е. подвижность носителей заряда численно равна скорости дрейфа в электрическом поле единичной напряженности.

С учетом равенства (1.21) выражение (1.9) для плотности тока примет вид:

$$\mathbf{J} = -en\mathbf{v} = en\mu\mathcal{E}, \quad (1.22)$$

так как вектор скорости электронов $\mathbf{v}$ направлен в противоположную сторону вектора $\mathcal{E}$.

Удельная проводимость на основании закона Ома может быть выражена при помощи (1.22) как

$$\sigma = J/\mathcal{E} = en\mu. \quad (1.23)$$

С учетом (1.21) удельная проводимость

$$\sigma = \frac{e^2 n \tau}{m}. \quad (1.24)$$

Если выразить τ из (1.4), то соотношение (1.24) можно записать в виде

$$\sigma = \frac{e^2 n l}{m v_0}. \quad (1.25)$$

В действительности же при определении τ надо учитывать полную скорость электрона, которая определяется тепловой скоростью v_0 и скоростью дрейфа v_d , т. е. $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_d$,

$$\tau = l/v. \quad (1.26)$$

Поскольку дрейфовая скорость электрона зависит от напряженности внешнего электрического поля, то его полная скорость $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_d$ и, следовательно, время свободного пробега τ являются функцией энергии частицы. Поэтому для нахождения дрейфовой скорости необходимо усреднять время свободного пробега с учетом функции распределения электронов по энергии. Это будет рассмотрено в гл. 6.

ГЛАВА ВТОРАЯ

ОСНОВЫ ЗОННОЙ ТЕОРИИ ПОЛУПРОВОДНИКОВ

2.1. УРАВНЕНИЕ ШРЕДИНГЕРА ДЛЯ КРИСТАЛЛА

Твердое тело, как известно, состоит из атомов, т. е. из ядер атомов и электронов. Ядра атомов образуют кристаллическую решетку, которая обладает свойством пространственной периодичности. При наложении внешнего электрического поля решетка практически не деформируется, хотя ядра атомов и заряжены. Это происходит потому, что силы, удерживающие ядра атомов в узлах кристаллической решетки, обычно значительно больше тех сил, которые создаются внешними электрическими полями. Те из элек-

тронов, которые не находятся близко к ядру атома и поэтому к нему не очень сильно притягиваются, могут передвигаться по твердому телу, создавая электрический ток. Однако при количественном описании этого явления возникают серьезные трудности. Они связаны с тем, что электроны являются заряженными частицами и при своем движении по твердому телу встречаются с другими электронами. Но так как между электронами действуют электрические силы отталкивания, то движение электрона оказывается зависимым от движения окружающих его электронов. Иными словами, в рассматриваемом случае надо решать не одноэлектронную, а многоэлектронную задачу. Поэтому для определения стационарных состояний и энергетического спектра совокупности большого числа атомных ядер и электронов в кристалле нужно решить уравнение Шредингера:

$$\hat{H}\Phi = \mathcal{E}\Phi, \quad (2.1)$$

где $\hat{H}$ — гамильтониан кристалла; Φ — его волновая функция; $\mathcal{E}$ — энергия кристалла.

Волновая функция кристалла зависит от координат всех электронов $\mathbf{r}_i$ и всех атомных ядер $\mathbf{R}_\alpha$:

$$\Phi = \Phi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N).$$

Оператор Гамильтона включает в себя:

1) оператор кинетической энергии электронов

$$\sum_i \left(-\frac{\hbar^2}{2m_0} \Delta_i \right),$$

где $\hbar$ — постоянная Планка h , деленная на 2π ($\hbar = h/2\pi$); m_0 — масса электрона; $\Delta_i = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}$ — оператор Лапласа для i -го электрона;

2) оператор кинетической энергии ядер

$$\sum_\alpha \left(-\frac{\hbar^2}{2M_\alpha} \Delta_\alpha \right),$$

где M_α — масса ядра; $\Delta_\alpha = \frac{\partial^2}{\partial X_\alpha^2} + \frac{\partial^2}{\partial Y_\alpha^2} + \frac{\partial^2}{\partial Z_\alpha^2}$;

3) потенциальную энергию попарного взаимодействия электронов

$$\frac{1}{2} \sum_i \sum_{j \neq i} \frac{e^2}{r_{ij}};$$

4) потенциальную энергию попарного взаимодействия ядер

$$V_0(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N);$$

5) потенциальную энергию взаимодействия электронов с ядрами

$$U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N).$$

С учетом этих составляющих гамильтониана уравнение Шредингера примет вид:

$$\left\{ \sum_i \left(-\frac{\hbar^2}{2m_0} \Delta_i \right) + \sum_{\alpha} \left(-\frac{\hbar^2}{2M_{\alpha}} \Delta_{\alpha} \right) + \frac{1}{2} \sum_{i \neq j} \sum_j \frac{e^2}{r_{ij}} + \right. \\ \left. + V_0(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) + U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \right\} \Phi = \mathcal{E} \Phi. \quad (2.2)$$

Число независимых переменных в (2.2) определяется полным числом частиц в кристалле, которые в 1 см^3 вещества составляют около 10^{23} . Такая задача в настоящее время не может быть решена в общем виде. Возможно лишь ее приближенное решение, которое достигается с помощью ряда последовательных приближений.

2.2. АДИАБАТИЧЕСКОЕ ПРИБЛИЖЕНИЕ И ВАЛЕНТНАЯ АППРОКСИМАЦИЯ

Разделим всю систему частиц на легкие (электроны) и тяжелые (атомные ядра). В равновесном состоянии средние значения кинетической энергии этих частиц одного порядка. Так как масса ядра намного больше массы электрона, т. е. $M_{\alpha} \gg m_0$, то скорости движения электронов намного превосходят скорости ядер (приблизительно на два порядка). При каждом изменении положения атомных ядер практически мгновенно устанавливается пространственное распределение электронов, соответствующее новому положению ядер. Поэтому можно считать, что движение электронов независимо от движения ядер, а значит, движение электронов и ядер происходит без обмена между ними энергией. Это позволяет в первом приближении рассматривать движение электронов в потенциальном поле фиксированных ядер. В этом случае волновая функция и энергия электронов будут некоторыми функциями, адиабатически меняющимися с изменением расположения ядер, координаты которых будут входить в эти функции как параметры. При изучении движения ядер, напротив, следует учитывать не мгновенное положение электронов, а поле, создаваемое их средним пространственным распределением. Такое приближенное рассмотрение называют адиабатическим, или приближением Борна — Оппенгеймера.

Воспользуемся наиболее грубым приближением, полагая, что ядра покоятся. Для этого случая координаты ядер $\mathbf{R}_1, \dots, \mathbf{R}_N$ уже не переменные, а параметры $\mathbf{R}_{10}, \dots, \mathbf{R}_{N0}$, представляющие собой координаты узлов кристаллической решетки. Для случая покоящихся ядер, когда $\mathbf{R}_{\alpha} = \mathbf{R}_{\alpha 0}$, кинетическая энергия ядер обращается в нуль, а энергия их взаимодействия V_0 принимает постоянное значение. Выбором начала отсчета энергии ее можно обратить

в нуль. С учетом этого уравнение (2.2) упрощается и теперь описывает движение электронов в поле покоящихся ядер. Для этого случая энергия электронов E_e и их волновая функция Ψ_e уже параметрически зависят от расположения ядер $R_{\alpha 0}$. В результате уравнение Шредингера для электронов, движущихся в поле покоящихся ядер, запишется в виде

$$\left\{ \sum_i \left(-\frac{\hbar^2}{2m_0} \Delta_i \right) + \frac{1}{2} \sum_i \sum_{i \neq j} \frac{e^2}{r_{ij}} + U(r_1, \dots, r_n, R_{10}, \dots, R_{N0}) \right\} \Psi_e = E_e \Psi_e. \quad (2.3)$$

Для дальнейшего сокращения числа переменных в уравнении (2.3) используется валентная аппроксимация. Принимается, что все электроны атома, кроме валентных, вместе с ядром образуют неподвижный атомный остаток (ион). При этом (2.3) записывается только для валентных электронов, движущихся в потенциальном поле фиксированных ионов.

Несмотря на проведенные упрощения, уравнение (2.3) не может быть решено в общем виде: по-прежнему мы имеем дело с задачей многих частиц, не поддающейся непосредственному решению.

2.3. ОДНОЭЛЕКТРОННОЕ ПРИБЛИЖЕНИЕ

Одним из наиболее распространенных методов решения многоэлектронной задачи для кристалла является метод Хартри—Фока, позволяющий многоэлектронную задачу свести к одноэлектронной. Его идея заключается в том, что *энергия попарного взаимодействия электронов заменяется взаимодействием каждого электрона с усредненным полем всех остальных электронов*. Пусть потенциальная энергия i -го электрона в том поле будет Ω_i . Она зависит от движения не только всех остальных электронов, но и от движения данного электрона, так как его движение оказывает влияние на движение остальных электронов. Поскольку поле определяет не только движение данного электрона, но и само зависит от его движения, то это поле получило название *самосогласованного*. Введение самосогласованного поля позволяет в уравнении (2.3) двойную сумму $\frac{1}{2} \sum_i \sum_{i \neq j} \frac{e^2}{r_{ij}}$ заменить суммой $\sum_i \Omega_i(r_i)$, каждый член кото-

рой зависит от координат одного электрона, т. е.

$$\frac{1}{2} \sum_i \sum_{i \neq j} \frac{e^2}{r_{ij}} = \sum_i \Omega_i(r_i), \quad (2.4)$$

где $\Omega_i(r_i)$ — потенциальная энергия i -го электрона в поле всех остальных электронов.

Аналогично потенциальную энергию взаимодействия электронов с ядрами можно также представить в виде суммы

$$U(\mathbf{r}_1, \mathbf{r}_2 \dots) = \sum_i U_i(\mathbf{r}_i), \quad (2.5)$$

где $U_i(\mathbf{r}_i)$ — потенциальная энергия i -го электрона в поле всех ядер.

С учетом (2.4) и (2.5) уравнение (2.3) преобразуется к виду

$$\left\{ \sum_i \left[-\frac{\hbar^2}{2m_0} \Delta_i + \Omega_i(\mathbf{r}_i) + U_i(\mathbf{r}_i) \right] \right\} \Psi_e = E_e \Psi_e. \quad (2.6)$$

Это уравнение можно записать как

$$\hat{H} \Psi_e = \left(\sum_i \hat{H}_i \right) \Psi_e = E_e \Psi_e, \quad (2.7)$$

где $\hat{H}_i$ — гамильтониан i -го электрона, равный:

$$\hat{H}_i = -\frac{\hbar^2}{2m_0} \Delta_i + \Omega_i(\mathbf{r}_i) + U_i(\mathbf{r}_i). \quad (2.8)$$

Из соотношения (2.7) следует, что гамильтониан кристалла представляет собой сумму гамильтонианов (2.8), каждый из которых зависит от координат одной частицы. Для этого случая волновая функция системы частиц может быть представлена как произведение волновых функций, описывающих состояние отдельных частиц системы:

$$\Psi_e(\mathbf{r}_1, \mathbf{r}_2 \dots) = \Psi_1(\mathbf{r}_1) \Psi_2(\mathbf{r}_2) \dots = \prod \Psi_i(\mathbf{r}_i). \quad (2.9)$$

Это означает, что электроны ведут себя независимо друг от друга (как бы не взаимодействуя), а полная энергия системы частиц равна сумме энергии отдельных электронов:

$$E_e = E_1 + E_2 + \dots = \sum_i E_i. \quad (2.10)$$

Таким образом, взведение самосогласованного поля позволяет рассматривать электроны в кристалле как невзаимодействующие частицы, т.е. движущиеся независимо друг от друга. Это и является основанием для представления электронов проводимости в виде идеального газа.

Обозначим через $\prod \Psi_{i(j)}$ произведение всех волновых функций, кроме Ψ_j . Тогда

$$[\prod \Psi_{i(1)}] \Psi_1 = [\prod \Psi_{i(2)}] \Psi_2 = \dots = \prod \Psi_i. \quad (2.11)$$

Учитывая, что оператор $\hat{H}_i$ действует только на волновую функцию Ψ_i , на основании (2.7), (2.9) и (2.11) получаем:

$$[\prod \Psi_{i(1)}] \hat{H}_1 \Psi_1 + [\prod \Psi_{i(2)}] \hat{H}_2 \Psi_2 + \dots = E_e \prod \Psi_i \quad (2.12)$$

Разделив обе части (2.12) на $\prod \Psi_i$ и учитывая соотношение (2.10) будем иметь:

$$\frac{1}{\Psi_1} \hat{H}_1 \Psi_1 + \frac{1}{\Psi_2} \hat{H}_2 \Psi_2 + \dots = E_1 + E_2 + \dots \quad (2.13)$$

Каждый член в левой части уравнения (2.13) зависит от координат лишь одного электрона, поэтому это уравнение эквивалентно системе следующих одноэлектронных уравнений:

$$\left. \begin{aligned} \hat{H}_1 \Psi_1 &= E_1 \Psi_1; \\ \hat{H}_2 \Psi_2 &= E_2 \Psi_2; \\ &\vdots \end{aligned} \right\} \quad (2.14)$$

Таким образом, введение самосогласованного поля позволяет задачу многих частиц свести к задаче для одного электрона:

$$\hat{H} \Psi = E \Psi, \quad (2.15)$$

где $\hat{H}$, $\Psi(\mathbf{r})$, E — соответственно гамильтониан, волновая функция и энергия электрона в кристалле.

Если ввести обозначение для потенциальной энергии электрона в кристалле через функцию $V(\mathbf{r})$, равную:

$$V(\mathbf{r}) = U(\mathbf{r}) + \Omega(\mathbf{r}), \quad (2.16)$$

то уравнение Шредингера для электрона кристалла запишется в виде

$$\left[-\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2.17)$$

Существуют различные способы рационального выбора вида функции $V(\mathbf{r})$ для решения одноэлектронной задачи. Наиболее часто для этого используются состояния электрона, находящегося в потенциальном поле всех ионов решетки, заряд которых в среднем скомпенсирован зарядом валентных электронов, т. е. в поле периодического потенциала. Иными словами, принимается, что член $V(\mathbf{r})$ в уравнениях (2.16) и (2.17) есть полный потенциал кристалла, обладающий трехмерной периодичностью решетки. Качественный вид $V(\mathbf{r})$ для одномерной решетки показан на рис. 2.1, где пунктирной кривой изображена потенциальная энергия изолированного атома $U_a(\mathbf{r})$.

В идеальном кристалле атомы периодически расположены в пространстве. Такое расположение атомов означает, что при смещении кристалла на вектор

$$\mathbf{a}_n = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3, \quad (2.18)$$

где $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ — периоды идентичности решетки по трем произвольным направлениям, а n_1 , n_2 , n_3 — произвольные целые числа, кристалл совмещается сам с собой. Из этого следует, что точка с ра-

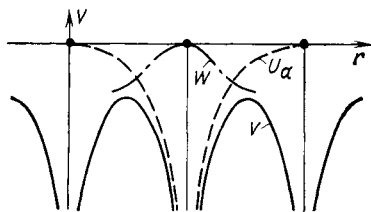


Рис. 2.1. Схематическое изображение периодического потенциала одномерной решетки. Сплошная кривая — потенциал решетки $V(\mathbf{r})$; пунктирная — атомный потенциал $U_a(\mathbf{r})$; штрихпунктирная — возмущающий потенциал $W(\mathbf{r})$

диусом-вектором $\mathbf{r}$ и точка с радиусом-вектором $\mathbf{r} + \mathbf{a}_n$ физически эквивалентны, поэтому

$$V(\mathbf{r} + \mathbf{a}_n) = V(\mathbf{r}). \quad (2.19)$$

Соотношение (2.19) выражает условие периодичности потенциального поля кристалла.

В периодическом поле кристалла, если собственное значение E невырождено, то волновая функция электрона $\Psi(\mathbf{r} + \mathbf{a}_n)$ невырождена и отличается от волновой функции $\Psi(\mathbf{r})$ только постоянным множителем, т. е.

$$\Psi(\mathbf{r} + \mathbf{a}_n) = C\Psi(\mathbf{r}). \quad (2.20)$$

Но из условия нормировки волновой функции следует, что

$$|C|^2 = 1, \quad (2.21)$$

поэтому можно положить, что

$$C = e^{ika_n}, \quad (2.22)$$

так как

$$|C|^2 = |e^{ika_n}|^2 = |\cos ka_n + i \sin ka_n|^2 = \cos^2 ka_n + \sin^2 ka_n = 1.$$

Здесь $\mathbf{k}$ — постоянный вектор, называемый волновым вектором, — характеризует квантовое состояние электрона в кристалле.

Поскольку произведение ka_n должно быть безразмерным, то волновой вектор имеет размерность обратной длины. Модуль волнового вектора, называемый волновым числом, выражает собой число длин волн, укладывающихся на отрезке 2π , а поэтому равен $k = 2\pi/\lambda$.

Из (2.20) и (2.22) следует, что

$$\Psi(\mathbf{r} + \mathbf{a}_n) = e^{ika_n}\Psi(\mathbf{r}), \quad (2.23)$$

откуда

$$\Psi(\mathbf{r}) = e^{-ika_n}\Psi(\mathbf{r} + \mathbf{a}_n) = e^{ikr}U_{\mathbf{k}}(\mathbf{r}), \quad (2.24)$$

где

$$U_{\mathbf{k}}(\mathbf{r}) = e^{-ik(\mathbf{r} + \mathbf{a}_n)}\Psi(\mathbf{r} + \mathbf{a}_n). \quad (2.25)$$

Функция $U_{\mathbf{k}}(\mathbf{r})$ обладает трехмерной периодичностью кристаллической решетки, так как согласно (2.23) и (2.25)

$$\begin{aligned} U_{\mathbf{k}}(\mathbf{r} + \mathbf{a}_l) &= e^{-ik(\mathbf{r} + \mathbf{a}_n + \mathbf{a}_l)}\Psi(\mathbf{r} + \mathbf{a}_n + \mathbf{a}_l) = \\ &= e^{-ik(\mathbf{r} + \mathbf{a}_n + \mathbf{a}_l)}e^{ika_l}\Psi(\mathbf{r} + \mathbf{a}_n) = e^{-ik(\mathbf{r} + \mathbf{a}_n)}\Psi(\mathbf{r} + \mathbf{a}_n) = U_{\mathbf{k}}(\mathbf{r}). \end{aligned} \quad (2.26)$$

Таким образом, стационарная волновая функция электрона в периодическом поле кристалла зависит от волнового вектора $\mathbf{k}$ и имеет вид:

$$\Psi_{\mathbf{k}}(\mathbf{r}) = e^{ikr}U_{\mathbf{k}}(\mathbf{r}), \quad (2.27)$$

где e^{ikr} — плоская волна, бегущая в направлении вектора $\mathbf{k}$; $U_{\mathbf{k}}(\mathbf{r})$ — некая функция координат, зависящая от волнового вектора $\mathbf{k}$ и имеющая периодичность решетки. Выражение (2.27) для $\Psi_{\mathbf{k}}(\mathbf{r})$ носит

название волны (или функции) Блоха. Следовательно, волновая функция электрона в периодическом поле кристалла представляет собой плоскую модулированную по амплитуде с периодичностью решетки кристалла волну, бегущую в направлении волнового вектора.

Если функцию Блоха (2.27) подставить в (2.15), то будем иметь:

$$\hat{H}\Psi_{\mathbf{k}}(\mathbf{r}) = E\Psi_{\mathbf{k}}(\mathbf{r}). \quad (2.28)$$

Из (2.28) следует, что энергия электрона в кристалле должна зависеть от волнового вектора $\mathbf{k}$, т. е. $E = E(\mathbf{k})$.

Следовательно, решением уравнения Шредингера для электрона в периодическом поле кристалла является бегущая плоская волна, модулированная с периодичностью решетки, а закон дисперсии для энергии электрона в кристалле зависит от волнового вектора $\mathbf{k}$.

Если бы полученные результаты на основе введения самосогласованного поля зависели от численного значения потенциала периодического поля $V(\mathbf{r})$, то в настоящее время, по-видимому, не существовало бы теории твердого тела, так как поля эти неизвестны и не могут быть определены ни теоретически, ни экспериментально. Однако для того чтобы получить фундаментальные результаты теории, нет необходимости знать численные значения силовых полей, достаточно лишь знать, что это поле периодически в пространстве и что его периоды совпадают с периодами решетки.

2.4. ПРИБЛИЖЕНИЕ СИЛЬНО СВЯЗАННЫХ ЭЛЕКТРОНОВ

При описании движения электрона в периодическом поле кристаллической решетки квантовая механика дает такие результаты, которые удобно сравнивать с квантовомеханическими результатами для изолированного атома. Известно, что спектр излучения свободного атома представляет собой набор дискретных линий. Как показывает квантовая теория, это объясняется тем, что электроны в изолированном атоме обладают дискретными значениями энергии.

Одно из положений физики гласит, что состояние устойчивого равновесия, в котором система может долго находиться, определяется минимумом потенциальной энергии. С этой точки зрения все электроны в атоме должны были бы находиться в наинизшем энергетическом состоянии, т. е. на энергетическом уровне, который расположен ближе всего к ядру. Однако квантовомеханические законы накладывают на это положение ограничение, сформулированное Паули. Согласно принципу Паули на одном и том же энергетическом уровне может находиться не более двух электронов. При этом электроны должны иметь противоположно направленные спины.

Рассмотрим теперь, что происходит с энергетическими уровнями при взаимодействии большого числа атомов, образующих кристалл. Уровни энергии внутренних электронов, расположенных ближе

к ядру, при этом почти не изменяются. Об этом можно судить по рентгеновским характеристическим спектрам, вид которых почти не зависит от соединения или агрегатного состояния вещества. Однако оптический спектр, обусловленный переходом самих внешних электронов, резко меняется.

Если считать, что кинетическая энергия электронов значительно больше пространственных изменений его потенциальной энергии, то периодический потенциал $V(\mathbf{r})$ можно рассматривать как малое возмущение свободного движения электронов. Такой подход, получивший название приближения почти свободных электронов, дает более или менее удовлетворительные результаты при решении некоторых задач для металлов.

Анализ физических свойств полупроводников более нагляден в приближении сильно связанных электронов, в котором считают, что состояние электрона в кристалле мало отличается от состояния его в изолированном атоме. Но такой подход применим только для электронов, находящихся на глубоких энергетических уровнях атомов, т. е. он применим для электронов, которые слабо взаимодействуют с атомами других узлов решетки. Поэтому приближения ни слабо, ни сильно связанных электронов не позволяют количественно описать состояние валентных электронов в кристалле. Другими словами, эти приближения не могут быть использованы для количественных расчетов энергетического спектра электронов конкретного вещества, но они хорошо иллюстрируют общие закономерности движения электрона в периодическом поле кристалла.

Пусть $\mathbf{r}$ — текущий радиус-вектор электрона, а $\mathbf{R}_g$ — радиус-вектор g -го узла кристаллической решетки, тогда $\mathbf{r} - \mathbf{R}_g$ есть расстояние между данной точкой $\mathbf{r}$ и ядром g , а волновая функция электрона в g -м атоме имеет вид:

$$\Psi_g = \Psi_g(\mathbf{r} - \mathbf{R}_g). \quad (2.29)$$

В приближении сильно связанных электронов считается, что состояние электрона в атоме кристалла мало меняется по сравнению с его состоянием в изолированном атоме. Поэтому волновую функцию электрона в кристалле можно представить в виде линейной комбинации атомных волновых функций (2.29):

$$\Psi(\mathbf{r}) = \sum_g a_g \Psi_g. \quad (2.30)$$

Чтобы выражение (2.30) для $\Psi(\mathbf{r})$ имело вид блоховской функции (2.27), необходимо положить:

$$a_g = e^{i\mathbf{k}\mathbf{R}_g}. \quad (2.31)$$

Тогда

$$\Psi(\mathbf{r}) = \sum_g e^{i\mathbf{k}\mathbf{R}_g} \Psi_g(\mathbf{r} - \mathbf{R}_g). \quad (2.32)$$

Выбор волновой функции электрона в кристалле в виде (2.32) вполне законен. Поскольку атомная волновая функция экспоненциально убывает, то вблизи g -го узла решетки

$$\Psi(\mathbf{r}) \approx \text{const } \Psi_g(\mathbf{r} - \mathbf{R}_g), \quad (2.33)$$

т. е. ведет себя как атомная волновая функция g -го узла.

Подставим выражение (2.30) в уравнение Шредингера (2.15) для электрона в кристалле. В результате получим:

$$\sum_g a_g (\hat{H} \Psi_g - E \Psi_g) = 0. \quad (2.34)$$

Если $V(\mathbf{r})$ — потенциальная энергия электрона в кристалле, то гамильтониан кристалла

$$\hat{H} = -\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}). \quad (2.35)$$

Для нахождения собственных значений энергии E электрона в кристалле при решении уравнения (2.34) будем считать, что энергетический спектр изолированного атома известен. Обозначим гамильтониан изолированного атома через

$$\hat{H}_a = -\frac{\hbar^2}{2m_0} \Delta + U_a(\mathbf{r}) \quad (2.36)$$

и запишем уравнение Шредингера для изолированного атома:

$$\hat{H}_a \Psi_a = E_a \Psi_a. \quad (2.37)$$

В этих формулах $U_a(\mathbf{r})$ — потенциальная энергия электрона в изолированном атоме; E_a — энергия некоторого уровня; Ψ_a — соответствующая уровню E_a волновая функция.

Вычитая выражение (2.36) из (2.35), будем иметь:

$$\hat{H} = \hat{H}_a + W(\mathbf{r}). \quad (2.38)$$

Здесь через

$$W(\mathbf{r}) = V(\mathbf{r}) - U_a(\mathbf{r}) \quad (2.39)$$

обозначена энергия возмущения для электрона в кристалле по сравнению с изолированным атомом (см. рис. 2.1).

Уравнение Шредингера (2.34) с учетом (2.38) примет вид:

$$\sum_g a_g \{ \hat{H}_g \Psi_g + W(\mathbf{r}) \Psi_g - E \Psi_g \} = 0. \quad (2.40)$$

Если учесть равенство (2.37), то это выражение преобразуется следующим образом:

$$\sum_g a_g \{ W(\mathbf{r}) \Psi_g + (E_a - E) \Psi_g \} = 0. \quad (2.41)$$

Умножим уравнение (2.41) на функцию Ψ_g^* , комплексно-сопряженную с Ψ_g , и проинтегрируем по всему объему кристалла τ :

$$\sum_g a_g \left\{ \int_{\tau} \Psi_g W(\mathbf{r}) \Psi_g^*, d\tau + (E_a - E) \int_{\tau} \Psi_g \Psi_g^*, d\tau \right\} = 0. \quad (2.42)$$

Введем обозначения:

$$\int_{\tau} \Psi_g W(\mathbf{r}) \Psi_g^* d\tau = A(\mathbf{R}_g - \mathbf{R}_{g'}), \quad (2.43)$$

Это есть обменный интеграл, зависящий от степени перекрытия волновых функций атомов и энергии возмущения $W(\mathbf{r})$. Он определяет обменную энергию взаимодействия.

Интеграл

$$\int_{\tau} \Psi_g \Psi_{g'}^* d\tau = S(\mathbf{R}_g - \mathbf{R}_{g'}) \quad (2.44)$$

зависит только от перекрытия атомных волновых функций.

С учетом (2.31), (2.43) и (2.44) уравнение (2.42) можно записать в виде

$$\sum_g A(\mathbf{R}_g - \mathbf{R}_{g'}) e^{i\mathbf{k}\mathbf{R}_g} + (E_a - E) \sum_g S(\mathbf{R}_g - \mathbf{R}_{g'}) e^{i\mathbf{k}\mathbf{R}_g} = 0. \quad (2.45)$$

Умножим полученное выражение на $e^{-i\mathbf{k}\mathbf{R}_{g'}}$. Это возможно, так как суммирование в (2.45) производится только по g , а g' неизменно. Тогда

$$\sum_g A(\mathbf{R}_g - \mathbf{R}_{g'}) e^{i\mathbf{k}(\mathbf{R}_g - \mathbf{R}_{g'})} + (E_a - E) \sum_g S(\mathbf{R}_g - \mathbf{R}_{g'}) e^{i\mathbf{k}(\mathbf{R}_g - \mathbf{R}_{g'})} = 0. \quad (2.46)$$

Упростим (2.46), введя вектор $\mathbf{q}$, соединяющий атомы g и g' :

$$\mathbf{q} = \mathbf{R}_g - \mathbf{R}_{g'}, \quad (2.47)$$

при этом получим:

$$\sum_{\mathbf{q}} A(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}} + (E_a - E) \sum_{\mathbf{q}} S(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}} = 0. \quad (2.48)$$

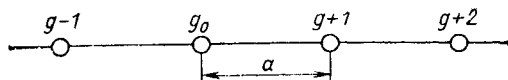
Отсюда можно найти собственные значения энергии электронов в кристалле:

$$E = E_a + \frac{\sum_{\mathbf{q}} A(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}}}{\sum_{\mathbf{q}} S(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}}}. \quad (2.49)$$

Из (2.49) следует, что энергия электрона в кристалле складывается из его энергии на соответствующем уровне в изолированном атоме и добавочного члена, представляющего собой периодическую функцию волнового вектора $\mathbf{k}$. Теперь вместо одного атомного уровня для электронов в кристалле имеем энергетическую зону, границы которой определяются минимальным и максимальным значениями второго члена в (2.49).

Рассмотрим более подробно добавочный член в (2.49). Для упрощения задачи будем считать, что атомные волновые функции очень быстро спадают и поэтому можно пренебречь их перекрытием даже для соседних атомов. Такое допущение является обоснованным в приближении сильно связанных электронов. При этом все члены в (2.44) обращаются в нуль, кроме члена, соответствующего $q = 0$,

Рис. 2.2. Схематическое изображение одномерной решетки



когда он равен единице из условия нормировки атомных волновых функций, т. е.

$$S(q) = \begin{cases} 1 & \text{при } q = 0; \\ 0 & \text{при } q \neq 0. \end{cases} \quad (2.50)$$

Отсюда следует, что

$$\sum_g S(q) e^{i k q} = 1. \quad (2.51)$$

Первый член суммы числителя дроби (2.49) для $q = 0$

$$A(0) = \int_{\tau} \Psi_0(\mathbf{r}) W(\mathbf{r}) \Psi_0^*(\mathbf{r}) d\tau = \int_{\tau} |\Psi_0(\mathbf{r})|^2 W(\mathbf{r}) d\tau = C \quad (2.52)$$

есть постоянная, не зависящая от $\mathbf{k}$.

Хотя мы пренебрегли перекрытием волновых функций соседних атомов, малое значение волновой функции атома вблизи нулевого узла частично компенсируется большой разностью $W(\mathbf{r}) = V(\mathbf{r}) - U_a(\mathbf{r})$ в этой области (см. рис. 2.1). Поэтому в сумме, состоящей в числителе дроби (2.49), следует учитывать значения обменных интегралов для узлов с $q \neq 0$, расположенных по соседству с нулевым узлом g_0 (рис. 2.2).

Если волновая функция соответствует s -состоянию, то эти обменные интегралы будут одинаковы для всех атомов. Пусть величина этого обменного интеграла равна A . Тогда с учетом (2.51) и (2.52) выражение (2.49) для энергии электрона в периодическом поле кристалла примет вид

$$E = E_0 + C + A \sum_g e^{i k q}. \quad (2.53)$$

В простой кубической решетке с периодом a каждый атом окружен шестью ближайшими атомами. Если оси x , y , z направить по ребрам куба, то около узла с $q = 0$ соседние узлы будут расположены на расстоянии $q = +a$ и $q = -a$. Поэтому

$$\begin{aligned} \sum_{g_0}^{1,6} e^{i k a} &= e^{i k_x a} + e^{-i k_x a} + e^{i k_y a} + e^{-i k_y a} + e^{i k_z a} + e^{-i k_z a} = \\ &= 2(\cos k_x a + \cos k_y a + \cos k_z a), \end{aligned} \quad (2.54)$$

так как $e^{i k a} = \cos k a + i \sin k a$, а $e^{-i k a} = \cos k a - i \sin k a$.

Выражение для энергии электрона в периодическом поле простой кубической решетки с учетом (2.54) примет вид:

$$E = E_a + C + 2A(\cos k_x a + \cos k_y a + \cos k_z a). \quad (2.55)$$

Из анализа данного выражения можно сделать ряд выводов относительно энергетического спектра электронов в кристаллах.

1. Уровень E_a изолированного атома при образовании кристаллической решетки в результате взаимодействия атомов смещается на величину C . Направление смещения уровня зависит от знака C .

2. Атомный уровень в кристаллической решетке расщепляется в полосу или зону, внутри которой энергия электрона периодически зависит от компонент волнового вектора k .

3. Экстремальные значения выражения (2.55), имеющие место при $\cos k_i a = \pm 1$ ($i = x, y, z$), будут:

$$E_{\max} = E_a + C + 6A,$$

$$E_{\min} = E_a + C - 6A.$$

Следовательно, для простой кубической решетки ширина энергетической зоны равна

$$E_{\max} - E_{\min} = 12 A, \quad (2.56)$$

т. е. зависит от величины обменного интеграла.

4. Каждый энергетический уровень изолированного атома в кристалле расщепляется в зону. Поскольку величина обменного интеграла определяется перекрытием электронных облаков соседних атомов, то, чем сильнее перекрываются волновые функции электронов соседних атомов, тем больше A и, следовательно, тем больше ширина энергетической зоны. В силу этого для более высоких атомных уровней из-за большего перекрытия волновых функций электронов образуется более широкая энергетическая зона (рис. 2.3).

5. Энергетические зоны в общем случае разделены запрещенными интервалами энергии E_g , называемыми запрещенными зонами (рис. 2.3).

6. С ростом энергии ширина энергетических зон увеличивается, а ширина запрещенных зон уменьшается (рис. 2.3).

7. Уровень E_a в изолированном атоме может быть вырожденным. В кристаллической решетке вырождение может быть частично или полностью снято. При этом атомный уровень расщепляется на несколько зон, число которых соответствует степени его вырождения.

Например, для p -состояния фактор вырождения $g=3$, так как $g = 2l + 1$, где l — азимутальное квантовое число, которое для p -состояния равно 1. Следовательно, из атомного p -состояния в кристалле возможно образование трех зон.

8. Энергия электрона в кристалле зависит от компонент волнового вектора k . Она является четной функцией волнового вектора k , т. е.

$$E(k) = E(-k).$$

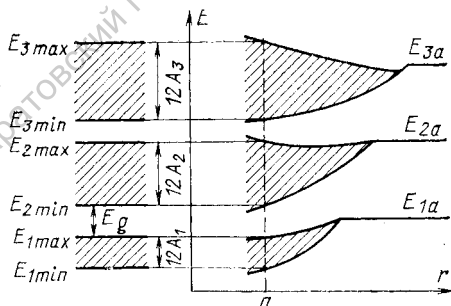


Рис. 2.3. Образование зон энергии из энергетических уровней при сближении атомов: a — постоянная решетки кристалла

9. При воздействии на кристалл температуры и давления, приводящих к изменению расстояния между атомами, будут изменяться область перекрытия волновых функций электронов и, следовательно, значение обменного интеграла. Это вызовет изменение ширины энергетических зон, в результате изменится и ширина запрещенной зоны между этими зонами.

10. Метод сильной связи неприменим к внешним валентным электронам атомов кристаллов, так как из-за большого перекрытия волновых функций электронов соседних атомов ширина энергетической зоны валентных электронов примерно равна расстоянию между уровнями энергии в изолированном атоме или превосходит их.

2.5. ЧИСЛО СОСТОЯНИЙ ЭЛЕКТРОНОВ В ЭНЕРГЕТИЧЕСКОЙ ЗОНЕ

Для подсчета числа состояний в энергетической зоне нужно при решении уравнения Шредингера для электрона учесть граничные условия при переходе от бесконечного идеального кристалла к ограниченному, т. е. учесть условия на краях кристалла. Любые условия на границах кристалла не отразятся существенным образом на состоянии электронов в его объеме, так как число узлов в любом реальном кристалле очень велико. Наиболее удобным для решения поставленной задачи является применение условия циклической граничности Борна—Кармана, обоснованием которого является периодичность свойств электрона в кристалле.

Рассмотрим кристалл в форме параллелепипеда с размерами по осям x , y , z соответственно L_x , L_y , L_z . Для кубической решетки с параметром a

$$L_x = aN_x; L_y = aN_y; L_z = aN_z, \quad (2.57)$$

где N_x , N_y , N_z — число атомов, укладываемых на соответствующих ребрах кристалла.

Потребуем, чтобы волновая функция Ψ имела на противоположных гранях параллелепипеда одно и то же значение. Такое условие не накладывает на вид волновой функции никаких физических ограничений, связанных с границами кристалла:

$$\Psi(x + L_x, y + L_y, z + L_z) = \Psi(x, y, z). \quad (2.58)$$

Уравнение (2.58) по существу выражает граничные условия циклической граничности Борна—Кармана.

Учитывая вид волновой функции для кристаллов (2.27), получаем:

$$\begin{aligned} \Psi(x + L_x, y + L_y, z + L_z) = \\ = U_k(\mathbf{r}) e^{ik_x L_x} e^{ik_y L_y} e^{ik_z L_z} = \Psi(x, y, z). \end{aligned} \quad (2.59)$$

Для выполнения условия (2.58) необходимо в выражении (2.59) принять:

$$e^{ik_x L_x} = 1; e^{ik_y L_y} = 1; e^{ik_z L_z} = 1. \quad (2.60)$$

Эти равенства выполнимы, если показатель экспоненты есть целое число, умноженное на $2\pi i$, т. е.

$$k_x L_x = 2\pi n_1; \quad k_y L_y = 2\pi n_2; \quad k_z L_z = 2\pi n_3, \quad (2.61)$$

где n_1, n_2, n_3 — произвольные целые числа. Отсюда следует, что

$$k_x = \frac{2\pi}{L_x} n_1; \quad k_y = \frac{2\pi}{L_y} n_2; \quad k_z = \frac{2\pi}{L_z} n_3, \quad (2.62)$$

где

$$n_1 = 0, \pm 1, \pm 2 \dots; \quad n_2 = 0, \pm 1, \pm 2 \dots; \\ n_3 = 0, \pm 1, \pm 2 \dots$$

Таким образом, компоненты волнового вектора k изменяются не непрерывно, а принимают ряд дискретных значений. В соответствии с этим оказывается квантованной и энергия электронов в разрешенной зоне.

С учетом значений волнового вектора k (2.62) можно записать волновую функцию (2.32) для одномерной задачи в виде

$$\Psi_k = \sum_g e^{ik_x a g} \Psi_g = \sum_g e^{i \frac{2\pi n_1 g}{N_x}} \Psi_g, \quad (2.63)$$

где учтено, что $R = ag$, g — номер атома, т. е. целое число, а $L_x = aN_x$.

Из выражения (2.63) следует, что волновая функция для $n_1 = \pm N_x$ будет Ψ_g и совпадает с функцией для $n_1 = 0$, а волновая функция для $n_1 = \pm (N_x + 1)$ совпадает с функцией для $n_1 = \pm 1$ и т. д. Это означает, что компонента k_x имеет N_x значений, соответствующих различным n_1 . При этом n_1 может принимать только значения $0, 1, 2, \dots, (N_x - 1)$, так как $E(k) = E(-k)$, т. е. n_1 изменяется только в пределах

$$0 \leq n_1 \leq N_x \text{ или } -N_x/2 \leq n_1 \leq N_x/2. \quad (2.64)$$

Из (2.57), (2.62) и (2.64) получаем, что компоненты вектора k находятся в следующих интервалах значений:

$$-\frac{\pi}{a} \leq k_x \leq \frac{\pi}{a}; \quad -\frac{\pi}{a} \leq k_y \leq \frac{\pi}{a}; \quad -\frac{\pi}{a} \leq k_z \leq \frac{\pi}{a}, \quad (2.65)$$

где k_x, k_y, k_z принимают соответственно N_x, N_y, N_z различных значений.

Следовательно, в разрешенной зоне кристалла имеется всего $N = N_x N_y N_z = L_x L_y L_z / a^3$ различных энергетических состояний (энергетических уровней), соответствующих различным k , равных числу элементарных ячеек в кристалле.

Согласно квантовой механике состояние электрона в атоме характеризуется главным квантовым числом n , азимутальным квантовым числом l , магнитным квантовым числом m и s_z — проекцией спина на ось z . Состояние электрона в кристалле согласно принципу Паули также должно описываться четырьмя квантовыми чи-

слами. Как следует из (2.62), тремя квантовыми числами являются проекции волнового вектора k_x , k_y , k_z , а четвертым квантовым числом должно быть s_z . Проекция s_z может принимать только два значения: $+1/2$ и $-1/2$. Это означает что в состоянии (k_x, k_y, k_z) может быть не более двух электронов. Но набор (k_x, k_y, k_z) согласно уравнению (2.55) определяет энергию $E(k)$ для данной зоны. Следовательно, на каждом энергетическом уровне зоны, который определяется волновым вектором k , в соответствии с принципом Паули может находиться не более двух электронов с противоположно направленными спинами. Таким образом, в простой энергетической зоне, возникшей из невырожденного атомного уровня, имеется $2N$ квантовых состояний, соответственно N энергетических уровней и не более $2N$ электронов. Если зона g кратно вырождена, то в ней может быть $2gN$ электронов. Из этого следует, что число квантовых состояний в зоне равно общему числу мест на уровнях изолированных атомов, из которых образовалась эта зона, т. е. имеет место сохранение числа состояний при образовании кристалла из атомов.

Оценим среднее расстояние между соседними уровнями энергии в разрешенной зоне. Примем значение параметра решетки $a = 0,4 \text{ нм} = 4 \cdot 10^{-8} \text{ см}$, что дает объем элементарной ячейки $a^3 = 64 \cdot 10^{-24} \text{ см}^3$. Количество элементарных ячеек в кристалле единичного объема $V = 1 \text{ см}^3$, равное числу состояний в зоне, составит:

$$N = V/a^3 = 1/64 \cdot 10^{-24} \approx 10^{22}.$$

При ширине зоны 1 эВ среднее расстояние между ее уровнями будет около 10^{-22} эВ , т. е. энергетическую зону можно считать квазинепрерывной.

2.6. КВАЗИИМПУЛЬС

Состояние электрона, свободно движущегося в пространстве, как известно, можно охарактеризовать энергией E и импульсом P . При этом связь между энергией и импульсом дается классической формулой

$$E = \frac{p^2}{2m_0}. \quad (2.66)$$

Согласно де Бройлю свободному электрону массы m_0 , движущемуся со скоростью v , соответствует волна, длина которой может быть определена из соотношения

$$\lambda = h/P = h/(m_0 v), \quad (2.67)$$

где h — постоянная Планка.

Так как волновое число — число волн, укладывающихся на длине $2\pi \text{ см}$, равно:

$$k = 2\pi/\lambda, \quad (2.68)$$

то импульс свободного электрона

$$P = \hbar k, \quad (2.69)$$

а его энергия

$$E = \frac{\hbar^2 k^2}{2m_0}, \quad (2.70)$$

где $\hbar = h/2\pi$ — квант действия.

Для электрона, движущегося в периодическом поле кристалла, можно ввести величину $\mathbf{p} = \hbar \mathbf{k}$, называемую квазиимпульсом. В соответствии с дискретным спектром k (2.62) квазиимпульс $\mathbf{p}$ также должен принимать ряд дискретных значений. Согласно (2.65) в кубической решетке квазиимпульс должен изменяться в пределах

$$-\frac{\pi \hbar}{a} \leq p_i \leq \frac{\pi \hbar}{a}; \quad i = x, y, z. \quad (2.71)$$

Если в (2.55) вместо k_i подставить значение p_i , найденное из (2.71), то для энергии электрона в кубическом кристалле получим:

$$E = E_a + C + 2A \left(\cos \frac{a}{\hbar} p_x + \cos \frac{a}{\hbar} p_y + \cos \frac{a}{\hbar} p_z \right). \quad (2.72)$$

Из соотношения (2.72) следует, что энергия электрона в кристалле — четная функция квазиимпульса, т. е.

$$E(\mathbf{p}) = E(-\mathbf{p}). \quad (2.73)$$

2.7. ЗОНЫ БРИЛЛЮЭНА

Как было показано в § 2.5, в энергетической зоне кристалла имеется N энергетических состояний, которым соответствуют значения компонент волнового вектора

$$k_i = \frac{2\pi}{L_i} n_j \quad (2.74)$$

и компонент квазиимпульса

$$p_i = \frac{2\pi \hbar}{L_i} n_j = \frac{h}{L_i} n_j, \quad (2.75)$$

где $i = x, y, z$, а $j = 1, 2, 3$.

Для кристалла с простой кубической решеткой согласно соотношениям (2.65) и (2.71) достаточно рассматривать изменение компонент k_i и p_i в пределах

$$\left. \begin{aligned} -\frac{\pi}{a} &\leq k_i \leq \frac{\pi}{a}; \\ -\frac{\pi \hbar}{a} &\leq p_i \leq \frac{\pi \hbar}{a}. \end{aligned} \right\} \quad (2.76)$$

Этим значениям квазиимпульса в системе координат (p_x, p_y, p_z) будет соответствовать некоторая область, построенная вокруг начала координат $p(000)$ и содержащая все возможные различные состояния. Эту область называют первой, или основной, зоной Бриллюэна,

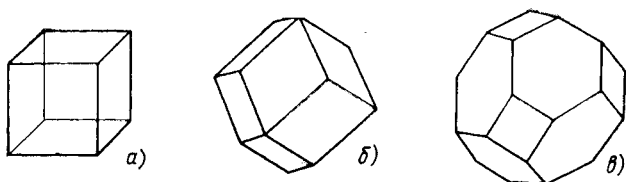


Рис. 2.4. Первая зона Бриллюэна для кристалла с простой кубической решеткой (а) кубической объемно-центрированной (б) и кубической грансцентрированной ршеткой (в)

Для кристалла с простой кубической решеткой первая зона Бриллюэна представляет собой куб (рис. 2.4, а) объемом

$$V_B^{(p)} = (2\pi\hbar/a)^3 = h^3/a^3. \quad (2.77)$$

В k -пространстве первая зона Бриллюэна для кристалла с простой кубической решеткой также является кубом, объем которого

$$V_B^{(k)} = (2\pi/a)^3 = 8\pi^3/a^3. \quad (2.78)$$

На рис. 2.5 схематично показано сечение первой зоны Бриллюэна для простой кубической решетки плоскостью $p_z = 0$. Крестиками отмечены точки, соответствующие различным состояниям электронов. Центр зоны Бриллюэна (начало координат) изображает состояние $n_1 = n_2 = n_3 = 0$, т. е. $p = 0$. Расстояние между ближайшими точками в зоне Бриллюэна согласно (2.75) равно $2\pi\hbar/L = h/L$. При этом, как следует из (2.63), точки, отмеченные на рисунке светлыми кружками, изображают состояния, которые физически эквивалентны соответствующим состояниям на противоположных гранях зоны Бриллюэна. Например, точки 1 и 2 изображают эквивалентные состояния, имеющие одну и ту же волновую функцию, и электроны, находящиеся в этих двух состояниях, будут иметь одинаковую энергию.

Первую зону Бриллюэна можно разбить на элементарные кубические ячейки объемом

$$\Delta V_B^{(p)} = (2\pi\hbar/L)^3 = h^3/a^3 N = h^3/V \quad (2.79)$$

и

$$\Delta V_B^{(k)} = (2\pi/L)^3 = 8\pi^3/a^3 N = 8\pi^3/V, \quad (2.79a)$$

где $V = L^3 = a^3 N_x N_y N_z = a^3 N$ — объем кристалла, а $N = N_x N_y N_z$ — полное число элементарных ячеек в кристалле.

Поскольку объем первой зоны Бриллюэна для кристалла с простой кубической решеткой равен $(h/a)^3$, а объем элементарной ячейки $h^3/a^3 N$, то число элементарных ячеек в ней составляет N , т. е. равно количеству энергетических состояний в зоне. Но в энергетической зоне может располагаться $2N$ электронов, следовательно, и в первой зоне Бриллюэна может быть $2N$ электронов, а в ее каждой элементарной ячейке может находиться только два электрона с антипараллельными спинами.

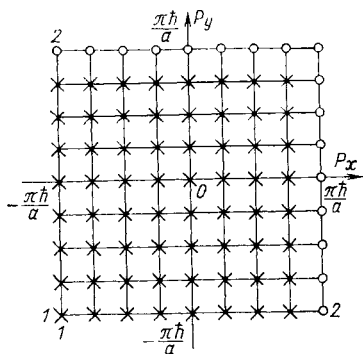


Рис. 2.5. Сечение первой зоны Бриллюэна для кристалла с простой кубической решеткой плоскостью $p_z = 0$

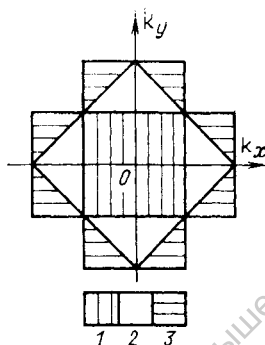


Рис. 2.6. Первые три зоны Бриллюэна для плоской квадратной решетки

Вторая и последующие зоны Бриллюэна, соответствующие второй и последующим энергетическим зонам, имеют более сложную конфигурацию, но их объем остается постоянным. Они также содержат N элементарных ячеек, каждой из которых можно сопоставить ячейку в первой зоне, изображающую эквивалентное состояние. В качестве примера на рис. 2.6 изображены в k -пространстве три первые бриллюэновские зоны для плоской квадратной решетки. Первая зона Бриллюэна для кристалла с простой кубической решеткой изображена на рис. 2.4, а, с кубической объемно-центрированной решеткой — на рис. 2.4, б и с кубической гранецентрированной решеткой на рис. 2.4, в.

2.8. ВОЗМОЖНОЕ ЗАПОЛНЕНИЕ ЭЛЕКТРОННЫХ СОСТОЯНИЙ ВАЛЕНТНОЙ ЗОНЫ

Рассмотрим несколько примеров возможного заполнения электронами квантовых состояний валентной зоны.

1. Зона заполнена электронами частично. В этом случае под действием внешнего электрического поля электроны переходят на соседние свободные энергетические уровни. Наличие свободных незанятых состояний в зоне дает возможность электронам двигаться в ней под действием внешнего электрического поля и переносить электрический заряд. Таким образом, для того чтобы в твердом теле протекал электрический ток, в валентной зоне должны быть свободные состояния.

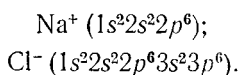
Типичными представителями веществ, обладающих хорошей проводимостью, являются одновалентные щелочные металлы Li, Na, K, Rb, Cs и благородные — Cu, Ag, Au. У щелочных элементов валентный электрон находится в ns -состоянии, а остальные электроны занимают состояния, которые они полностью заполняют.

В качестве примера рассмотрим натрий, у которого 11 электронов распределены следующим образом: $\text{Na}^{(11)} (1s^2 2s^2 2p^6 3s)$. При образовании кристалла уровни атомов натрия расщепляются в зоны. Зоны, образованные из уровней $1s$, $2s$ и $2p$, полностью заполнены, так как в их $2N$ -, $2N$ - и $6N$ -состояниях находятся $2N$, $2N$ и $6N$ электронов. В зоне валентных электронов, образованной из уровня $3s$, в $2N$ -состояниях находятся N электронов, которые заполняют энергетическую зону только наполовину.

2. В валентной зоне все возможные электронные состояния заняты, но эта зона перекрывается со свободной зоной, не занятой электронами. В этом случае электроны валентной зоны будут переходить на энергетические уровни свободной зоны и осуществлять перенос электрического заряда. Этот случай реализуется у элементов второй группы. Щелочноземельные элементы имеют два валентных электрона, находящихся в ns -состоянии. Например, у $\text{Mg}^{(12)} (1s^2 2s^2 2p^6 3s^2) 3s$ - и $3p$ -зоны перекрываются, в результате чего электроны верхних уровней $3s$ -зоны занимают нижние уровни $3p$ -зоны.

3. Число возможных состояний в валентной зоне равно количеству валентных электронов атомов, образовавших кристалл. В этом случае все состояния в зоне заняты электронами, на каждом уровне зоны располагается по два электрода с противоположно направленными спинами. Поэтому внешнее электрическое поле не может создать направленного движения такой совокупности электронов, ибо в заполненной зоне электроны могут только взаимно обмениваться местами. Следовательно, такой кристалл не может проводить ток, он является диэлектриком.

В качестве примера рассмотрим энергетический спектр кристалла хлористого натрия. При образовании кристалла NaCl валентный электрон натрия переходит к атому хлора. В итоге возникают отрицательный ион хлора и положительный ион натрия



У иона хлора в $3p$ -состоянии имеется шесть электронов, что соответствует наиболее устойчивому энергетическому состоянию, т. е. все квантовые состояния заняты и на каждом уровне находится по два электрона с антипараллельными спинами. Из этого состояния в кристалле образуется зона, в которой все $6N$ -состояний заняты электронами. Это будет валентная зона. У иона натрия в $3s$ -состоянии не имеется электрона, а поэтому из этого уровня в кристалле образуется зона, в которой все $2N$ квантовых состояний свободны. Это будет зона проводимости.

Проанализируем энергетический спектр кристаллов, образованных из элементов IV группы таблицы Менделеева, обладающих кристаллической решеткой типа алмаза. В нее входят углерод (алмаз),

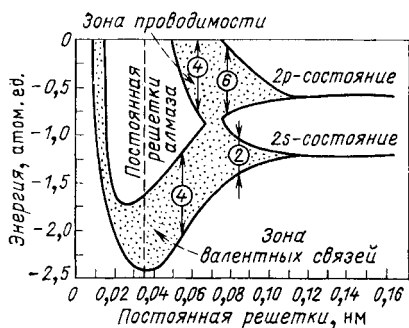


Рис. 2.7. Образование энергетических зон из атомных уровней в алмазе

кремний, германий и серое олово. Электронная структура этих атомов такова (см. рис. 1.4 и 1.5), что в твердом состоянии у них в образовании ковалентной связи принимают участие четыре электрона каждого атома. При этом, как следует из рис. 2.7, зоны, образованные из ns - и np -состояний, перекрываются, образуя общую зону с числом состояний $8N$. С уменьшением межатомного расстояния эта зона затем расщепляется на две зоны с $4N$ квантовыми состояниями в каж-

дой. Нижняя зона содержит $4N$ заполненных электронами состояний — это валентная зона, а у верхней зоны $4N$ квантовых состояний свободны — это зона проводимости.

2.9. ЗАВИСИМОСТЬ ЭНЕРГИИ ЭЛЕКТРОНА ОТ ВОЛНОВОГО ВЕКТОРА У ДНА И ПОТОЛКА ЭНЕРГЕТИЧЕСКОЙ ЗОНЫ

Энергия электрона в простой кубической решетке, которая выражается уравнением (2.55), определяется величиной и знаком обменного интеграла. В свою очередь знак обменного интеграла A , как следует из (2.43), зависит от степени перекрытия атомных волновых функций соответствующего состояния и знака энергии возбуждения $W(\mathbf{r})$. Энергия возмущения $W(\mathbf{r})$ согласно рис. 2.1 всегда является отрицательной. Поэтому для s -состояния, имеющего сферическую электронную оболочку, $A < 0$. Если учесть знак обменного интеграла, то энергия электрона в зоне, образованной из s -состояния, будет равна:

$$E = E' - 2|A|(\cos k_x a + \cos k_y a + \cos k_z a), \quad (2.80)$$

где $E' = E_a + C$.

Эта зависимость для значений волнового вектора k_x , меняющегося в пределах первой зоны Бриллюэна, изображена на рис. 2.8. Кривая $E(k)$ обращена вогнутостью вверх для s -зоны и для компонент волновых векторов k_y и k_z .

На рис. 2.8 точка $k_1 = 0$ определяет состояние с минимальной энергией. Разрешенные уровни энергии соответствуют значениям $k = \frac{2\pi}{L_x} n$. Точка $k_N = \frac{2\pi}{L_x} (N - 1)$ отвечает состоянию с максимальным значением энергии.

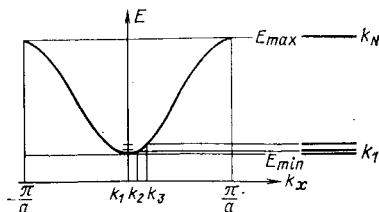


Рис. 2.8. Зависимость $E(k_x)$ для зоны, образованной из s -состояния

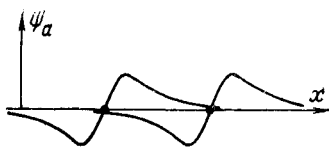


Рис. 2.9. Атомные функции p -состояния, описываемые уравнением (2.81)

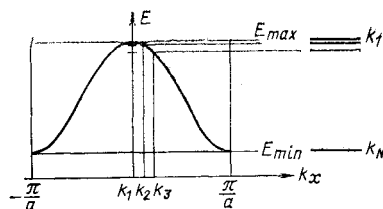


Рис. 2.10. Зависимость $E(k_x)$ для зоны, образованной из p -состояния

Для p -состояния, атомные функции которого имеют вид:

$$\frac{x}{r} f(r), \quad (2.81)$$

перекрывающиеся части функций Ψ_a для атомов с различными координатами, как видно из рис. 2.9, имеют в основном противоположные знаки.

Как следует из рис. 2.9, произведение перекрывающихся частей функций одинакового знака мало, поэтому можно считать, что $A > 0$; тогда вместо (2.80) получаем:

$$E = E'' + 2|A|(\cos k_x a + \cos k_y a + \cos k_z a). \quad (2.82)$$

Зависимость энергии от k_x для зоны, образованной из p -состояния, представлена на рис. 2.10. Здесь при $k_1 = 0$ имеется максимум энергии, а k_N соответствует минимуму энергии.

Для зоны, образованной из d -состояния, зависимость $E(k)$ аналогична зависимости для зоны из s -состояния и график $E(k)$ имеет вид, подобный изображенному на рис. 2.8.

Рассмотрим более детально (2.55) для энергии электронов, движущихся в потенциальном поле простой кубической решетки. Если выбрать начало отсчета энергии так, чтобы $E_a + C = 0$, то энергия электронов в энергетической зоне, образованной из s -состояния, запишется в виде

$$E = -2A(\cos k_x a + \cos k_y a + \cos k_z a), \quad (2.83)$$

где $A > 0$.

Обозначим:

$$E_{min} = -6A = E_d; \quad (2.84)$$

$$E_{max} = 6A = E_s. \quad (2.85)$$

Проанализируем (2.83) в окрестности точки $\mathbf{k} = 0$, т. е. вблизи центра первой зоны Бриллюэна. Разложим функцию $E(\mathbf{k})$ в ряд и ограничимся первыми членами, приняв во внимание, что $k_i a \ll 1$:

$$\begin{aligned} E &= -2A \left\{ \left[1 - \frac{(k_x a)^2}{2} \right] + \left[1 - \frac{(k_y a)^2}{2} \right] + \left[1 - \frac{(k_z a)^2}{2} \right] \right\} = \\ &= -6A + Aa^2(k_x^2 + k_y^2 + k_z^2) = -6A + Aa^2 k^2. \end{aligned} \quad (2.86)$$

Таким образом, энергия электрона у дна энергетической зоны не зависит от направления волнового вектора $\mathbf{k}$ и пропорциональна k^2 .

По аналогии с (2.70) для энергии свободного электрона запишем E для электрона в кристалле, используя (2.84) и (2.86), в виде

$$E = E_d + \frac{\hbar^2}{2m_n^*(d)} k^2, \quad (2.87)$$

где

$$m_n^*(d) = \hbar^2 / 2Aa^2 \quad (2.88)$$

есть эффективная масса электрона у дна энергетической зоны.

Рассмотрим теперь формулу (2.83) для состояний, близких к верхнему краю зоны. Введем вектор $\mathbf{k}'$ с компонентами $k'_i = \frac{\pi}{a} - k_i$. При этом у потолка зоны $k'_i a \ll 1$ и разложение (2.83) в ряд с учетом соотношения (2.85) примет вид:

$$\begin{aligned} E &= -2A [\cos(\pi - k'_x a) + \cos(\pi - k'_y a) + \cos(\pi - k'_z a)] = \\ &= 2A \left\{ \left[1 - \frac{(k'_x a)^2}{2} \right] + \left[1 - \frac{(k'_y a)^2}{2} \right] + \left[1 - \frac{(k'_z a)^2}{2} \right] \right\} = \\ &= 6A - Aa^2 (k')^2 = E_v + \frac{\hbar^2}{2m_n^*(v)} (k')^2. \end{aligned} \quad (2.89)$$

Итак, энергия электрона у верхнего края энергетической зоны также пропорциональна $(k')^2$ и не зависит от направления $\mathbf{k}'$, но эффективная масса в этом случае оказывается отрицательной

$$m_n^*(v) = -\frac{\hbar^2}{2Aa^2}. \quad (2.90)$$

Поскольку масса по самой своей сути не может быть отрицательна, то для состояний у верхнего края энергетической зоны вводится понятие дырки, представляющей собой частицу с положительным зарядом, равным заряду электрона, и положительной эффективной массой:

$$m_p^*(v) = -m_n^*(v) = \frac{\hbar^2}{2Aa^2}. \quad (2.91)$$

Следовательно, эффективная масса дырки у верхнего края энергетической зоны совпадает с эффективной массой электрона у дна этой зоны и обратно пропорциональна значению обменного интеграла. Но обменный интеграл согласно соотношению (2.56) определяет ширину энергетической зоны. Отсюда следует, что эффективная масса носителей заряда обратно пропорциональна ширине энергетической зоны. Это означает, что чем выше зона энергии, тем она шире и тем меньше в ней эффективная масса носителей заряда.

На рис. 2.11 показана зависимость $E(\mathbf{k})$ для трех соседних зон одномерной решетки, образованных из s -, p - и d -состояний атомов для области изменения k в пределах $\pm\pi/a$. Каждая разрешенная зона содержит N энергетических уровней, среднее расстояние между которыми примерно равно 10^{-22} эВ. В точках $k = 0$ и $k = \pm\pi/a$ энергия электрона принимает попеременно минимальное и максимальное значения.

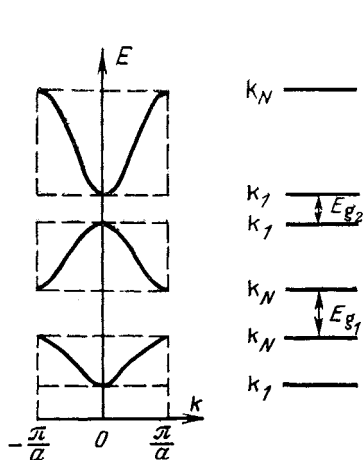


Рис. 2.11. Зависимость $E(k)$ для трех соседних зон, образованных из s -, p - и d -состояний атомов одномерной решетки

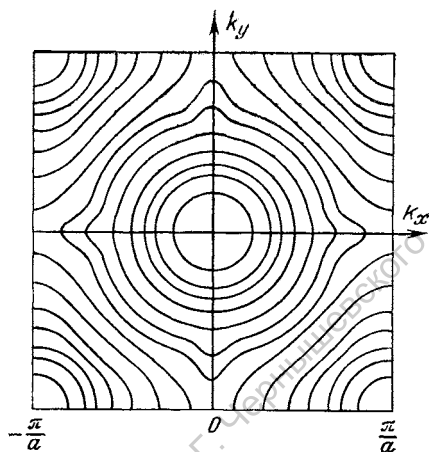


Рис. 2.12. Изоэнергетические линии для квадратной плоской решетки

Если рассматривать пространство k_x, k_y, k_z или p_x, p_y, p_z , то уравнение

$$E(\mathbf{k}) = \text{const} \text{ или } E(\mathbf{p}) = \text{const} \quad (2.92)$$

определяет поверхность постоянной энергии или изоэнергетическую поверхность. Согласно (2.86) и (2.89) энергия электрона у дна и потолка энергетической зоны не зависит от направления вектора $\mathbf{k}$ и изоэнергетические поверхности вблизи центра зоны Бриллюэна при $k = 0$ и около точек $k = \pm\pi/a$ являются сферами. По мере удаления от этих точек сферы начинают деформироваться.

Вид изоэнергетических поверхностей (кривых) для квадратной плоской решетки приведен на рис. 2.12. В центре и вблизи углов квадрата изоэнергетические кривые представляют собой окружности, которые искажаются по мере удаления от этих точек.

2.10. ДВИЖЕНИЕ ЭЛЕКТРОНОВ В КРИСТАЛЛЕ ПОД ДЕЙСТВИЕМ ВНЕШНЕГО ЭЛЕКТРИЧЕСКОГО ПОЛЯ

Как было установлено, энергетический спектр электронов в периодическом поле кристалла имеет вид квазинепрерывных зон, разделенных промежутками запрещенных значений энергии (см. рис. 2.3).

Рассмотрим теперь движение электронов в кристалле, находящемся во внешнем электрическом поле, направленном вдоль оси x , $\mathcal{E} = \mathcal{E}_x = \text{const}$. В этом случае на электрон, находящийся в периодическом поле кристалла, будет действовать внешняя сила $\mathbf{F}$. Будем считать, что сила $\mathbf{F}$ достаточно мала и не может вызвать переход электрона из одной энергетической зоны в другую, т. е.

$$Fa \ll E_g, \quad (2.93)$$

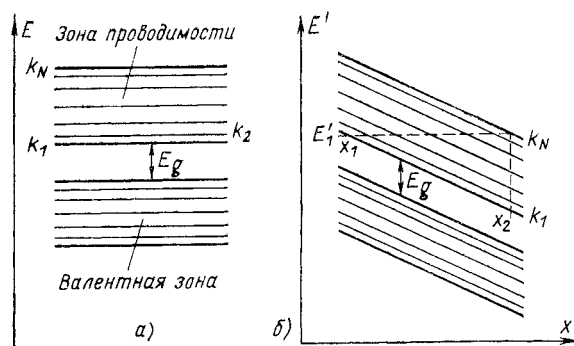


Рис. 2.13. Наклон энергетических зон во внешнем электрическом поле

где a — постоянная решетки; E_g — ширина соответствующей запрещенной зоны.

Потенциальная энергия электрона во внешнем электрическом поле напряженности $\mathcal{E}$ в данном случае будет определяться его координатой x , т. е.

$$U(r) = -e\mathcal{E}x. \quad (2.94)$$

Уравнение Шредингера для этого случая имеет вид:

$$\left\{ -\frac{\hbar^2}{2m_0} \Delta + [V(r) + U(r)] \right\} \Psi = E' \Psi. \quad (2.95)$$

Для упрощения задачи определения собственных значений энергии E' из (2.95) примем, что потенциал внешнего поля U мало изменяется в интервале изменения r , равном постоянной решетки, так что в области каждого атома можно считать $U = \text{const}$. При этом получаем уравнение Шредингера, подобное уравнению для стационарного случая:

$$\left[-\frac{\hbar^2}{2m_0} \Delta + V(r) \right] \Psi = (E' - U) \Psi = E \Psi, \quad (2.96)$$

где E — энергия электрона в отсутствие внешнего электрического поля.

Отсюда следует, что энергия электрона в кристалле при наличии внешнего электрического поля равна:

$$E' = E + U(r). \quad (2.97)$$

Тогда вместо (2.55) с учетом (2.94) можно написать:

$$E' = E_a + C + 2A(\cos k_x a + \cos k_y a + \cos k_z a) - e\mathcal{E}x. \quad (2.98)$$

Из (2.98) следует, что если задавать все возможные значения энергии E' , то получим график зависимости положения энергетических зон от координаты (рис. 2.13, б). Этот график можно интерпретировать как наклон энергетических зон под действием внешнего электрического поля.

Можно показать, что ширина зоны для каждого значения x осталась прежней. Действительно, для верхнего и нижнего уровней зоны имеем:

$$\left. \begin{aligned} E_{max} &= E_a + C + 6A - e\mathcal{E}x; \\ E_{min} &= E_a + C - 6A - e\mathcal{E}x. \end{aligned} \right\} \quad (2.99)$$

Разность $E_{max} - E_{min} = 12A$ не зависит от x и совпадает с шириной зоны для стационарного случая.

В отсутствие внешнего электрического поля электрон находится на одном из уровней энергетической зоны и обладает определенным значением волнового вектора k , например k_2 на рис. 2.13, а. Это означает, что в стационарном состоянии электрон перемещается вдоль кристалла на любые расстояния, имея постоянные значения квазиимпульса и волнового вектора.

При движении электрона в кристалле под действием внешнего электрического поля по закону сохранения энергии его полная энергия остается постоянной, т. е.

$$E' = E + U = \text{const}, \quad (2.100)$$

но изменяется соотношение между его кинетической и потенциальной энергиями.

На рис. 2.13, б значение полной энергии электрона кристалла при наличии внешнего электрического поля показано пунктирной прямой, параллельной оси абсцисс. Но в этом случае электрон при движении по кристаллу переходит с одного уровня энергетической зоны на ее другие уровни, в результате чего имеет место изменение волнового вектора электрона от k_1 до k_N , а следовательно, и его квазиимпульса от p_1 до p_N . При этом энергия $E(k)$ электрона меняется от E_{min} до E_{max} . Теперь электрон, как следует из рис. 2.13, б, перемещается не по всему кристаллу, а только в пределах от x_1 до x_2 , т. е. в области $\Delta x = x_2 - x_1$, равной:

$$\Delta x = \frac{E_{max} - E_{min}}{e\mathcal{E}} = \frac{12A}{e\mathcal{E}}. \quad (2.101)$$

Таким образом, во внешнем однородном электрическом поле зоны энергии идеального кристалла наклоняются, при этом уровни энергии поднимаются, если $U > 0$, и опускаются, если $U < 0$. В этом случае происходит перераспределение электронов, так как они движутся в некоторой области Δx , ширина которой обратно пропорциональна напряженности внешнего электрического поля. При движении в этой области в энергетической зоне электрон переходит с уровня на уровень и у электрона имеет место изменение волнового вектора и квазиимпульса, а следовательно, имеет место зависимость их от времени, т. е.

$$k = k(t); \quad p = p(t). \quad (2.102)$$

Найдем закон изменения квазиимпульса и волнового вектора от времени, т. е. закон, который описывает движение электрона в кристалле при наличии внешнего электрического поля.

Как известно из квантовой механики, движение свободного электрона с волновым вектором $\mathbf{k}$ можно описать с помощью волнового пакета, представляющего собой суперпозицию плоских волн с непрерывно меняющимися значениями k в пределах $2\Delta k$ (от $k - \Delta k$ до $k + \Delta k$).

Волновая функция волнового пакета для одномерного случая имеет вид

$$\Psi(x, t) = \int_{k-\Delta k}^{k+\Delta k} a(k) e^{i(kx - \omega t)} dk, \quad (2.103)$$

где ω — циклическая частота.

Движение волнового пакета характеризуется групповой скоростью $v_{\text{гп}} = \frac{d\omega}{dk}$, которая равна скорости перемещения какой-либо точки пакета, например его максимума. Координату этого максимума можно найти из условия $\bar{x} - \frac{d\omega}{dk} t = 0$. Отсюда следует, что

$$v = \frac{\bar{x}}{t} = \frac{d\omega}{dk} = v_{\text{гп}}, \quad (2.104)$$

т. е. средняя скорость движения свободного электрона v равна групповой скорости волнового пакета

$$v = v_{\text{гп}} = \frac{d\omega}{dk}. \quad (2.105)$$

Если воспользоваться соотношением для энергии $E = \hbar\omega$, то средняя скорость движения свободного электрона будет определяться выражением вида

$$v = \frac{d\omega}{dk} = \frac{1}{\hbar} \frac{dE}{dk} = \frac{dE}{dP}, \quad (2.106)$$

где $P = \hbar k$ — импульс свободного электрона.

Движение электрона в кристалле описывается волновой функцией (2.27), которая определяется набором атомных волновых функций с разным значением k . Поскольку $k = \frac{2\pi}{aN} n$, где $n = 0, 1, \dots, (N-1)$, $a \approx 10^{-8}$ см, $N \approx 10^{22}$ см $^{-3}$ и $aN \approx 10^{14}$, то волновую функцию Ψ можно рассматривать как совокупность плоских волн, для которых k меняется почти непрерывно. В силу этого движение электрона в кристалле можно охарактеризовать волновым пакетом, составленным из блоховских функций. Поэтому (2.106) будет справедливо и для средней скорости движения электрона в кристалле

$$\langle v \rangle = v = \frac{1}{\hbar} \frac{dE}{dk} = \frac{dE}{dp} \quad (2.107)$$

или для трехмерного случая

$$\mathbf{v} = \frac{1}{\hbar} \text{grad}_k E(\mathbf{k}) = \frac{1}{\hbar} \frac{dE}{d\mathbf{k}} = \frac{dE}{d\mathbf{p}}, \quad (2.108)$$

где $\mathbf{p} = \hbar \mathbf{k}$ — квазиимпульс.

Таким образом, *средняя скорость движения электрона в кристалле определяется производной энергии по квазиимпульсу.*

Вычислим среднюю квантовомеханическую скорость электрона в простой кубической решетке в приближении сильной связи. Из уравнений (2.83) и (2.108) получаем:

$$\begin{aligned} \mathbf{v} &= \frac{2aA}{\hbar} (i \sin k_x a + j \sin k_y a + k \sin k_z a) = \\ &= \frac{2aA}{\hbar} \left(i \sin \frac{a}{\hbar} p_x + j \sin \frac{a}{\hbar} p_y + \right. \\ &\quad \left. + k \sin \frac{a}{\hbar} p_z \right), \end{aligned} \quad (2.109)$$

где i, j, k — единичные векторы.

Из (2.109) следует, что *средняя скорость движения электронов в кристалле зависит как от абсолютной величины, так и от направления волнового вектора $\mathbf{k}$ (или квазиимпульса $\mathbf{p}$).*

Для одномерного движения, полагая $k_y = k_z = 0$ и $k_x = k$, имеем:

$$E = -2A \cos ka; \quad (2.110)$$

$$v = \frac{2aA}{\hbar} \sin ka. \quad (2.111)$$

Эти зависимости представлены на рис. 2.14, а и б. Из рисунков следует, что на краях зоны ($k = \pm\pi/a$) $E(\mathbf{k})$ экстремальна, т. е. $\text{grad}_{\mathbf{k}} E(\mathbf{k}) = 0$. Следовательно, в точках экстремума энергии у нижнего и верхнего краев энергетической зоны скорость движения электрона равна нулю. В точках же внутри зоны Бриллюэна электроны имеют наибольшие значения скорости. Средняя квантовомеханическая скорость электрона по всей зоне

$$\langle v \rangle = \int_{-\pi/a}^{\pi/a} v(k) dk = \frac{2aA}{\hbar} \int_{-\pi/a}^{\pi/a} \sin ka dk = 0. \quad (2.112)$$

Таким образом, *в отсутствие внешнего электрического поля средняя скорость движения электрона в кристалле по всей энергетической зоне равна нулю.*

Рассмотрим случай, когда на электрон в кристалле действует внешняя сила $\mathbf{F}$. Пусть $E(\mathbf{k})$ — энергия электрона в зоне, в которой он движется, а его скорость $\mathbf{v}$. Тогда согласно закону сохранения энергии имеем для одномерного движения:

$$\frac{dE}{dt} = Fv, \quad (2.113)$$

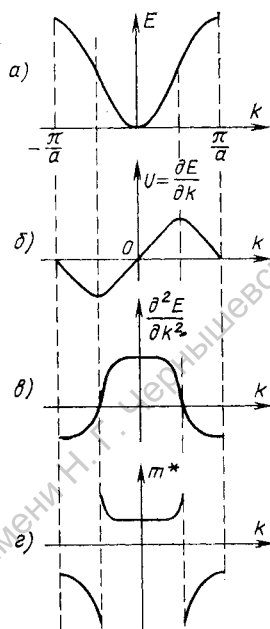


Рис. 2.14. Зависимость энергии (а), скорости (б), величины d^2E/dk^2 (в) и эффективной массы (г) от волнового вектора для кубической решетки

т. е. скорость изменения энергии частицы равна работе внешней силы в единицу времени.

Так как

$$\frac{dE}{dt} = \frac{dE}{dp} \frac{dp}{dt}, \quad (2.114)$$

то из сравнения равенств (2.113) и (2.114) с учетом (2.107) следует, что

$$\frac{dp}{dt} = F \quad \text{или} \quad \frac{d\mathbf{p}}{dt} = \mathbf{F}. \quad (2.115)$$

Это уравнение по форме совпадает с уравнением движения Ньютона: скорость изменения квазиимпульса равна приложенной силе. Следовательно, в периодическом поле кристалла $V(\mathbf{r})$ квазиимпульс электрона $\mathbf{p} = \hbar \mathbf{k}$ в уравнении движения (2.115) выполняет роль импульса свободного электрона.

Пусть внешняя сила $\mathbf{F}$ не зависит от времени. В этом случае движение электрона в $\mathbf{p}$ -пространстве будет описываться уравнением

$$\mathbf{p}(t) = \mathbf{p}_0 + \mathbf{F}t \quad (2.116)$$

и траекторией движения будет прямая линия, направление которой задается направлением внешней силы $\mathbf{F}$. Но электрон, двигаясь в $\mathbf{p}$ -пространстве (или $\mathbf{k}$ -пространстве), переходит от одной изоэнергетической поверхности к другой (см. рис. 2.12). Следовательно, внешняя сила $\mathbf{F}$ меняет не только квазиимпульс, но также и энергию электрона.

Рассмотрим теперь, как меняется импульс $\mathbf{P}$ электрона кристалла в отсутствие внешнего электрического поля. В кристалле с идеальной структурой, имеющей строго периодическое поле, электрон движется, оставаясь на одном и том же уровне энергетической зоны. Поскольку квазиимпульс электрона постоянен, то, значит, $\frac{d\mathbf{p}}{dt} = 0$. Но со стороны поля решетки на электрон кристалла действует сила $\mathbf{F}_{\text{кр}}$, она и определяет изменение его импульса $\mathbf{P}$, т. е.

$$\frac{d\mathbf{P}}{dt} = \mathbf{F}_{\text{кр}} \quad (2.117)$$

Так как $\mathbf{F}_{\text{кр}}$ является периодической функцией с периодом решетки, то и импульс электрона, движущегося в кристалле, не остается постоянным, а периодически меняется.

При наличии внешнего электрического поля на электрон действует как внешняя сила $\mathbf{F}$, так и сила $\mathbf{F}_{\text{кр}}$, обусловленная действием поля решетки. Поэтому изменение импульса электрона во времени будет определяться силой, действующей со стороны поля кристалла, и внешней силой, т. е.

$$\frac{d\mathbf{P}}{dt} = \mathbf{F}_{\text{кр}} + \mathbf{F}. \quad (2.118)$$

Итак, если структура кристаллической решетки идеальна, то ее потенциальное поле $V(\mathbf{r})$ строго периодическое. В периодическом

поле решетки электрон движется вдоль всего кристалла, имея постоянный квазиимпульс, а значит имея постоянную энергию. В этом случае электрон, перемещаясь по кристаллу, остается на одном и том же уровне энергетической зоны. Но поскольку $p = \text{const}$, а значит, $k = \text{const}$, то согласно (2.111) скорость движения электрона постоянна: $v = \text{const}$, следовательно, его ускорение $a = dv/dt = 0$. Это значит, что в периодическом поле решетки электрон движется без ускорения. Другими словами, в строго периодическом поле решетки электрон движется как свободная частица, без сопротивления, не рассеиваясь.

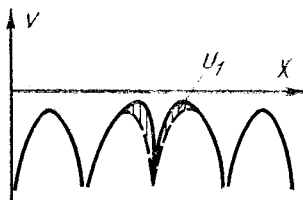


Рис. 2.15. Схематическое изображение искаженного потенциала одномерной решетки

Если кристалл с идеальной структурой поместить во внешнее электрическое поле, то, как следует из (2.115), движение электрона будет подобно движению свободной частицы под действием внешней силы F . Только теперь электрон, перемещаясь в ограниченной области кристалла Δx (2.101), в энергетической зоне переходит с уровня на уровень и его энергия изменяется.

Рассмотрим теперь случай, когда поле кристаллической решетки не является строго периодическим. Пусть поле решетки подобно изображенному на рис. 2.15. Для участка кристалла, где нарушена идеальность структуры, поле решетки можно представить как поле периодическое $V(r)$, на которое наложено некоторое дополнительное силовое поле U_1 , не обладающее периодичностью решетки. Но согласно (2.115) под действием неперiodичной части потенциального поля U_1 квазиимпульс электрона должен измениться. Из этого следует, что при любых нарушениях идеальной периодичности потенциального поля решетки происходит изменение квазиимпульса электрона. Это значит, что на любых нарушениях идеальности структуры решетки происходит рассеяние электронных волн.

В реальных кристаллах потенциальное поле не является строго периодическим. Оно искажено тепловыми колебаниями атомов решетки и наличием в ней разного типа дефектов. Рассеяние электронных волн на этих неоднородностях структуры решетки и является физической причиной, обуславливающей сопротивление вещества при прохождении через него электрического тока.

2.11. ЭФФЕКТИВНАЯ МАССА НОСИТЕЛЕЙ ЗАРЯДА

Допустим, что свободный электрон с массой m_0 находится в однородном электрическом поле с напряженностью $\mathcal{E}$. На электрон будет действовать сила $F = -e\mathcal{E}$, под воздействием которой электрон приобретает ускорение

$$a = \frac{1}{m_0} F = -\frac{1}{m_0} e\mathcal{E}, \quad (2.119)$$

направленное так же, как и внешняя сила, т. е. против поля с напряженностью $\mathbf{E}$.

Для электрона в кристалле, находящегося во внешнем электрическом поле, учитывая (2.107) и (2.115), можно записать:

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial t} = \frac{\partial^2 E}{\partial \mathbf{p}^2} \mathbf{F}. \quad (2.120)$$

Обобщая (2.120) для трехмерного случая, получаем:

$$\left. \begin{aligned} a_x &= \frac{\partial^2 E}{\partial p_x^2} F_x + \frac{\partial^2 E}{\partial p_x \partial p_y} F_y + \frac{\partial^2 E}{\partial p_x \partial p_z} F_z; \\ a_y &= \frac{\partial^2 E}{\partial p_y \partial p_x} F_x + \frac{\partial^2 E}{\partial p_y^2} F_y + \frac{\partial^2 E}{\partial p_y \partial p_z} F_z; \\ a_z &= \frac{\partial^2 E}{\partial p_z \partial p_x} F_x + \frac{\partial^2 E}{\partial p_z \partial p_y} F_y + \frac{\partial^2 E}{\partial p_z^2} F_z. \end{aligned} \right\} \quad (2.121)$$

В этом случае вектор ускорения $\mathbf{a}$ не совпадает по направлению с вектором внешней силы $\mathbf{F}$.

Совокупность величин $\frac{\partial^2 E}{\partial p_i \partial p_j} = \frac{1}{m_{ij}}$, связывающих векторы $\mathbf{a}$ и $\mathbf{F}$, является тензором второго ранга:

$$\mathbf{m}^{-1} = \left\| \begin{array}{ccc} \frac{\partial^2 E}{\partial p_x^2} & \frac{\partial^2 E}{\partial p_x \partial p_y} & \frac{\partial^2 E}{\partial p_x \partial p_z} \\ \frac{\partial^2 E}{\partial p_y \partial p_x} & \frac{\partial^2 E}{\partial p_y^2} & \frac{\partial^2 E}{\partial p_y \partial p_z} \\ \frac{\partial^2 E}{\partial p_z \partial p_x} & \frac{\partial^2 E}{\partial p_z \partial p_y} & \frac{\partial^2 E}{\partial p_z^2} \end{array} \right\| = \left\| \begin{array}{ccc} m_{xx}^{-1} m_{xy}^{-1} m_{xz}^{-1} \\ m_{yx}^{-1} m_{yy}^{-1} m_{yz}^{-1} \\ m_{zx}^{-1} m_{zy}^{-1} m_{zz}^{-1} \end{array} \right\|. \quad (2.122)$$

Величины $\frac{\partial^2 E}{\partial p_i \partial p_j} = m_{ij}^{-1}$ называют компонентами или элементами тензора. Поскольку размерность квазиимпульса совпадает с размерностью импульса, то размерность компонент тензора m_{ij}^{-1} есть размерность обратной массы, а размерность $m_{ij} = \left(\frac{\partial^2 E}{\partial p_i \partial p_j} \right)^{-1}$ есть размерность массы. Поэтому по аналогии с выражением (2.119) для свободного электрона тензор (2.122) называют тензором обратной эффективной массы.

Запишем выражение для энергии электрона в кристалле в малой окрестности экстремальной точки $\mathbf{k}_0$. Для этого разложим $E(\mathbf{k})$ в окрестности точки $\mathbf{k}_0$ в ряд по степеням компонент разности $\mathbf{k} - \mathbf{k}_0$ и ограничимся членами второго порядка малости по $\mathbf{k} - \mathbf{k}_0$. Поскольку точка $\mathbf{k}_0$ соответствует экстремуму энергии $E(\mathbf{k})$, то, учитывая, что для экстремума $\frac{\partial E}{\partial \mathbf{k}} = 0$, будем иметь:

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{1}{2} \sum_{ij} \left[\left(\frac{\partial^2 E}{\partial k_i \partial k_j} \right)_{\mathbf{k}_0} (k_i - k_{0i}) (k_j - k_{0j}) \right]. \quad (2.123)$$

Здесь i, j независимо пробегает значения x, y, z .

Учитывая, что

$$\frac{1}{\hbar^2} \left(\frac{\partial^2 E}{\partial k_i \partial k_j} \right)_{\mathbf{k}_0} = \left(\frac{\partial^2 E}{\partial p_i \partial p_j} \right)_{\mathbf{p}_0} = \frac{1}{m_{ij}}, \quad (2.124)$$

получаем:

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{\hbar^2}{2} \sum_{ij} \frac{1}{m_{ij}} (k_i - k_{0i})(k_j - k_{0j}) \quad (2.125)$$

или

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{1}{2} \sum_{ij} \frac{1}{m_{ij}} (p_i - p_{0i})(p_j - p_{0j}). \quad (2.126)$$

Скорость движения электрона в кристалле в малой окрестности экстремальной точки $\mathbf{k}_0$ будет равна:

$$v_i = \frac{\partial E}{\partial p_i} = \sum_j \frac{p_j - p_{0j}}{m_{ij}}, \quad (2.127)$$

где $i = x, y, z$ и $j = x, y, z$, например $v_x = \frac{p_x - p_{0x}}{m_{xx}} + \frac{p_y - p_{0y}}{m_{xy}} + \frac{p_z - p_{0z}}{m_{xz}}$. В векторной форме

$$\mathbf{v} = \frac{1}{\mathbf{m}} (\mathbf{p} - \mathbf{p}_0). \quad (2.128)$$

Следовательно, в общем случае скорость движения электрона в кристалле равна скалярному произведению квазиимпульса на тензор обратной эффективной массы.

Тензор (2.122) симметричен относительно главной диагонали, т. е. $m_{ij}^{-1} = m_{ji}^{-1}$, так как $\frac{\partial^2 E}{\partial p_i \partial p_j} = \frac{\partial^2 E}{\partial p_j \partial p_i}$. Выбрав соответствующую систему координат, можно свести симметричный тензор к диагональному виду, когда отличны от нуля только члены, расположенные на главной диагонали:

$$\mathbf{m}^{-1} = \begin{vmatrix} m_{xx}^{-1} & 0 & 0 \\ 0 & m_{yy}^{-1} & 0 \\ 0 & 0 & m_{zz}^{-1} \end{vmatrix} = \begin{vmatrix} m_1^{-1} & 0 & 0 \\ 0 & m_2^{-1} & 0 \\ 0 & 0 & m_3^{-1} \end{vmatrix}, \quad (2.129)$$

где $m_1^{-1} = m_{xx}^{-1}$; $m_2^{-1} = m_{yy}^{-1}$; $m_3^{-1} = m_{zz}^{-1}$.

Тогда тензором, обратным тензору обратной эффективной массы, будет тензор эффективной массы

$$\{\mathbf{m}^{-1}\}^{-1} = \mathbf{m} = \begin{vmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{vmatrix}. \quad (2.130)$$

Значения m_i ($i = 1, 2, 3$) называют компонентами тензора эффективной массы.

Если учесть компоненты тензора эффективной массы (2.130), то энергия электрона вблизи экстремальной точки $\mathbf{k}_0$ запишется в виде

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{\hbar^2 (k_x - k_{0x})^2}{2m_1} + \frac{\hbar^2 (k_y - k_{0y})^2}{2m_2} + \frac{\hbar^2 (k_z - k_{0z})^2}{2m_3} \quad (2.131)$$

или

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{(p_x - p_{0x})^2}{2m_1} + \frac{(p_y - p_{0y})^2}{2m_2} + \frac{(p_z - p_{0z})^2}{2m_3}, \quad (2.132)$$

а скорость

$$v_i = \frac{p_i - p_{0i}}{m_i} = \frac{\partial E}{\partial p_i}. \quad (2.133)$$

Из (2.132) следует, что поверхности равной энергии в $\mathbf{p}$ -пространстве $E(\mathbf{p}) - E(\mathbf{p}_0) = \text{const}$ вблизи экстремума энергии представляют собой эллипсоиды, а m_1, m_2, m_3 — эффективные массы вдоль основных осей эллипсоидальных поверхностей энергии.

Запишем уравнение эллипсоида в канонической форме

$$\frac{(p_x - p_{0x})^2}{a^2} + \frac{(p_y - p_{0y})^2}{b^2} + \frac{(p_z - p_{0z})^2}{c^2} = 1, \quad (2.134)$$

где полуоси эллипсоида a, b, c , направленные по координатным осям x, y, z , определяются из соотношений:

$$\left. \begin{aligned} a^2 &= 2m_1 [E(\mathbf{p}) - E(\mathbf{p}_0)]; \\ b^2 &= 2m_2 [E(\mathbf{p}) - E(\mathbf{p}_0)]; \\ c^2 &= 2m_3 [E(\mathbf{p}) - E(\mathbf{p}_0)]. \end{aligned} \right\} \quad (2.135)$$

Как известно, градиент направлен по нормали к поверхности уровня. Поскольку скорость движения электрона есть градиент энергии в пространстве квазиимпульса (2.133), следовательно, она направлена по нормали к изоэнергетической поверхности. Поэтому для эллипсоидальных поверхностей энергии (2.134) в общем случае направления векторов скорости $\mathbf{v}$ и квазиимпульса $\mathbf{p}$ не совпадают (рис. 2.16 и 2.17, а). Они будут совпадать по направлению только вдоль осей эллипсоидов. При этом, если принять $E(\mathbf{p}_0) = E_0$, то

$$p_i - p_{0i} = \sqrt{2m_i(E - E_0)}; \quad (2.136)$$

$$v = \frac{p_i - p_{0i}}{m_i} = \sqrt{\frac{2(E - E_0)}{m_i}}. \quad (2.137)$$

Из соотношения (2.137) следует, что при одной и той же энергии скорость движения вдоль осей эллипсоида обратно пропорциональна корню из соответствующей компоненты эффективной массы. Но компонента эффективной массы согласно (2.135) определяет размер оси эллипсоида. Поэтому, чем сильнее вытянут эллипсоид, т. е. чем больше m_i , тем меньше скорость движения электрона в этом направлении.

Если симметрия кристалла такова, что две главные оси тензора физически эквивалентны (например, x и y), соответствующие компоненты эффективной массы равны, т. е. $m_1 = m_2$, а следовательно, $a = b$. Уравнение изоэнергетической поверхности в этом случае упрощается:

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{(p_x - p_{0x})^2 + (p_y - p_{0y})^2}{2m_1} + \frac{(p_z - p_{0z})^2}{2m_3}. \quad (2.138)$$

Рис. 2.16. Скорость и квазимпульс электрона на эллипсоидальной поверхности энергии. Вектор $\mathbf{v}$ нормален к поверхности энергии; векторы $\mathbf{v}$ и $\mathbf{p}$ в общем случае неколлинеарны

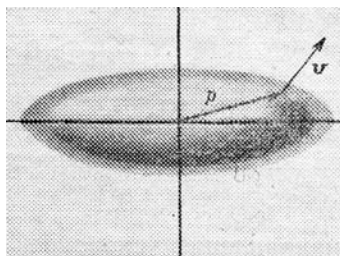
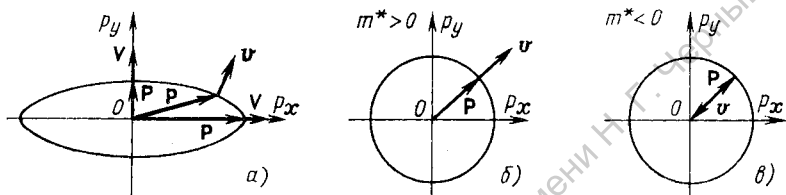


Рис. 2.17. Направления радиуса-вектора и нормали к изоэнергетической поверхности:

a — эллипсоидальные поверхности энергии; b — сферические поверхности энергии, $m^* > 0$; θ — сферические поверхности, $m^* < 0$.



Изоэнергетическая поверхность, описываемая уравнением (2.138), является эллипсоидом вращения, у которого ось вращения совпадает с осью p_z . Так как m_z — эффективная масса вдоль оси симметрии, т. е. вдоль оси вращения эллипсоида, то она получила название продольной эффективной массы и обозначается m_l . Величина $m_l = m_2$ — эффективная масса поперек осей симметрии, а поэтому ее называют поперечной эффективной массой и обозначают m_t . Если $m_l < m_z$, то эллипсоид вращения вытянут вдоль оси вращения, при этом он вытянут тем больше, чем больше отношение эффективных масс m_z/m_l . Если $m_l > m_z$, то эллипсоид сжат вдоль оси вращения.

Для кристаллов, обладающих кубической симметрией, все главные оси эквивалентны. При этом $m_l = m_2 = m_z = m^*$ и тензор эффективной массы вырождается в скаляр (тензор нулевого ранга). В этом случае изоэнергетические поверхности описываются уравнением

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{\mathbf{p}^2}{2m^*} = \text{const} \quad (2.139)$$

и представляют собой сферы (рис. 2.17, б, в). Выражение для эффективной массы имеет вид:

$$\frac{1}{m^*} = \frac{\partial^2 E}{\partial p^2} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial k^2}. \quad (2.140)$$

Следовательно, для кристалла кубической симметрии эффективная масса носителей заряда обратно пропорциональна кривизне изоэнергетических поверхностей в $\mathbf{k}$ -пространстве (или $\mathbf{p}$ -пространстве).

стве). В том случае, когда электрон находится в окрестности минимума энергии, т. е. в окрестности дна зоны проводимости (рис. 2.14, а)

$$\frac{\partial^2 E}{\partial k^2} > 0 \text{ и } m^* = \text{const} > 0, \quad (2.141)$$

т. е. электроны ведут себя как отрицательно заряженные частицы с положительной эффективной массой (рис. 2.14, з).

При этом согласно выражениям (2.120) и (2.140) получаем:

$$\mathbf{F} = m^* \mathbf{a}, \quad (2.142)$$

т. е. направление ускорения совпадает с направлением внешней силы, а значит, противоположно направлению внешнего электрического поля $\mathcal{E}$.

Так как ускорение $a = \frac{dv}{dt}$, а $F = \frac{dp}{dt}$, то квазиимпульс электрона в кубическом кристалле равен:

$$\mathbf{p} = m^* \mathbf{v}, \quad (2.143)$$

т. е. скорость движения электрона совпадает по направлению с квазиимпульсом (рис. 2.17, б).

Следовательно, под действием внешнего электрического поля движение электрона, находящегося у дна энергетической зоны кубического кристалла, подобно движению свободной частицы, масса которой равна m^* .

Из (2.142) следует, что ускорение электрону в кристалле сообщает только внешняя сила. Действие электрического поля решетки проявляется в том, что при наличии внешней силы движение электрона определяется не его обычной массой m_0 , а эффективной массой m^* .

В окрестности максимума энергии, т. е. в окрестности вершины валентной зоны (рис. 2.14, а),

$$\frac{\partial^2 E}{\partial k^2} < 0 \text{ и } m^* = \text{const} < 0 \quad (2.144)$$

(рис. 2.14, з), а, следовательно, скорость направлена против квазиимпульса (рис. 2.17, в). Из (2.144) следует, что направление ускорения электрона, находящегося у верхнего края зоны, противоположно направлению действующей на него внешней силы:

$$\mathbf{a} = - \frac{\mathbf{F}}{m^*}. \quad (2.145)$$

Но в однородном электрическом поле на электрон действует сила, равная:

$$\mathbf{F} = - e\mathcal{E}, \quad (2.146)$$

поэтому для электрона, находящегося вблизи потолка энергетической зоны, у которого $m^* < 0$, ускорение, сообщаемое внешним электрическим полем, будет:

$$\mathbf{a} = \frac{-|e|\mathcal{E}}{-|m^*|} = \frac{e\mathcal{E}}{m^*}, \quad (2.147)$$

т. е. ускорение движения носителя заряда направлено по полю $\mathcal{E}$,

Такой носитель заряда в окрестности вершины валентной зоны ведет себя как частица с положительным зарядом и положительной эффективной массой и носит название дырки.

По мере удаления от краев зоны зависимость $E(\mathbf{k})$ отклоняется от квадратичного закона. В точках перегиба $\mathbf{k} = \pm \pi/2a$ кривой $E(\mathbf{k})$ (рис. 2.14)

$$\frac{\partial^2 E}{\partial k^2} = 0 \text{ и } m^* \rightarrow \infty.$$

Из этого следует, что применение соотношения (2.115) для описания движения носителей заряда в кристалле под действием внешнего электрического поля возможно только для носителей заряда, находящихся вблизи дна или потолка энергетической зоны.

На практике наиболее важными оказываются два случая заполнения энергетической зоны.

1. Зона почти свободна. В этом случае количество электронов n в зоне много меньше числа возможных квантовых состояний $2N$. Электроны, стремясь занять низшие энергетические уровни, располагаются у дна зоны. При этом $m_n^* = \text{const} > 0$.

2. Зона почти заполнена электронами. В такой зоне имеется некоторое количество свободных квантовых состояний. Состояния, свободные от электронов, находятся в самой верхней части зоны, т. е. дырки располагаются у потолка зоны, где их энергия минимальна. Следовательно, энергия дырок отсчитывается в направлении, противоположном отсчету энергии для электрона. В этом случае $m_p^* = \text{const} > 0$.

При температуре абсолютного нуля валентная зона полупроводника заполнена электронами, а зона проводимости свободная. При повышении температуры небольшая часть электронов из валентной зоны в результате теплового возбуждения переходит в зону проводимости. При этом реализуется рассмотренный выше случай — наличие небольшого количества электронов у дна зоны проводимости и дырок у верхнего края валентной зоны. Так как ширина зоны проводимости больше ширины валентной зоны, а значение эффективной массы обратно пропорционально ширине энергетической зоны, эффективная масса дырки должна быть больше эффективной массы электрона, т. е.

$$m_p^* > m_n^*.$$

Следовательно, дырка и электрон проводимости не только отличаются знаком своего заряда, но имеют и разные величины эффективных масс.

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 $\mathbf{B} = \text{const}$. На

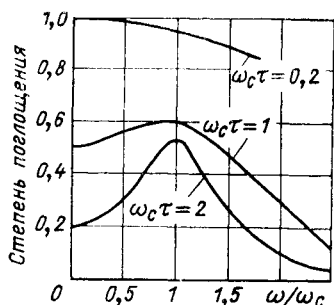


Рис. 2.18. Поглощение при циклотронном резонансе в зависимости от напряженности постоянного магнитного поля (в единицах ω/ω_c)

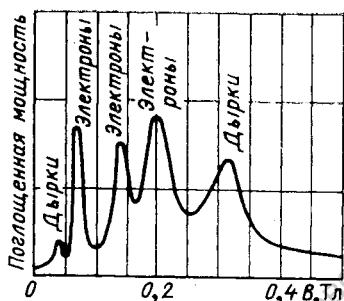


Рис. 2.19. Поглощение при циклотронном резонансе в германии (B — индукция)

электрон, движущийся со скоростью v , будет действовать сила Лоренца

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

Под действием этой силы электрон будет двигаться по окружности в плоскости, перпендикулярной магнитному полю. Круговая частота вращения ω_c , которая носит название **циклотронной**, определяется из равенства центробежной силы и силы Лоренца

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

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.

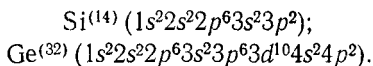
2.13. ЗОННАЯ СТРУКТУРА НЕКОТОРЫХ ПОЛУПРОВОДНИКОВ

В качестве примера рассмотрим зонную структуру кремния и германия, а также арсенида галлия и антимонида индия — полупроводников типа $A^{III}B^V$.

Кремний и германий имеют кристаллическую структуру типа алмаза, представляющую собой две гранцентрированные кубические решетки, сдвинутые одна относительно другой на $1/4$ пространственной диагонали. Элементарная ячейка содержит два атома. Для них первая зона Бриллюэна не является кубом, а имеет форму четырнадцатигранника, изображенного на рис. 2.20, где значения составляющих волнового вектора (k_x , k_y , k_z) даны в единицах $2\pi/a$ (a — ребро куба решетки).

Расчет зонной структуры полупроводников представляет собой сложную задачу. Для германия и кремния зависимость $E(k)$ для произвольно выбранной точки k является уравнением 146-й степени. Поэтому энергия была вычислена лишь для некоторых симметрично расположенных точек зоны Бриллюэна, для которых расчет значительно упрощается. Но и при этом пришлось численно решать уравнение 16-й степени. Данные для промежуточных точек в зоне Бриллюэна были получены интерполированием. Результаты теоретических расчетов уточнялись сравнением с данными экспериментов, в частности по циклотронному резонансу.

У атома кремния имеется 14 электронов, а у атома германия — 32 электрона, которые распределены по состояниям следующим образом:



У них последняя оболочка не заполнена, в ней в p -состоянии имеется два электрона с параллельными спинами. Поскольку зона проводимости и валентная зона кремния и германия включают p -состояние

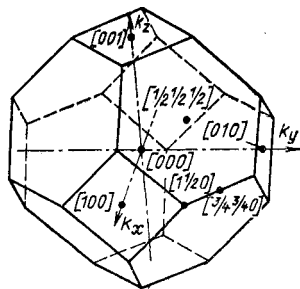


Рис. 2.20. Первая зона Бриллюэна полупроводника типа алмаза

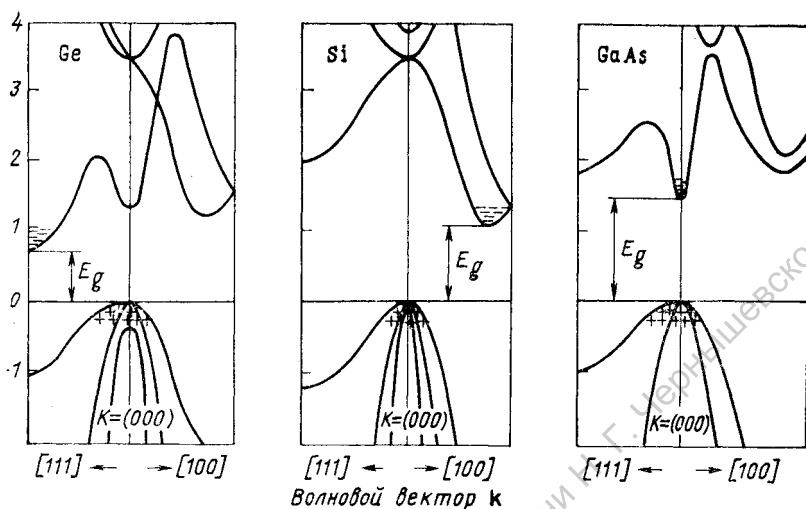


Рис. 2.21. Энергетическая зонная структура германия, кремния и арсенида галлия E_g — ширина запрещенной зоны; знаком плюс обозначены дырки в валентных зонах; знаком минус — электроны в зонах проводимости

(см. рис. 2.7), для которого в кристалле вырождение снимается, то каждая из них представляет собой наложение трех различных зон. На рис. 2.21 они представлены тремя ветвями $E(k)$. Эта зависимость неодинакова для разных кристаллографических направлений.

Одна из ветвей $E(k)$ зоны проводимости как у кремния, так и у германия (рис. 2.21) лежит значительно ниже других. Положения абсолютно минимума энергии определяет дно зоны проводимости. Минимумы энергии называют также д о л и н а м и.

Абсолютный минимум зоны проводимости у германия расположен в направлении осей $[111]$ (рис. 2.21), поэтому имеется восемь эквивалентных минимумов энергии, координаты которых в единицах $2\pi/a$ имеют вид $(1/2, 1/2, 1/2)$ и лежат на границе зоны Бриллюэна (на рис. 2.20 они показаны жирной точкой). Изоэнергетические поверхности около абсолютных минимумов энергии (или долин) представляют собой эллипсоиды вращения относительно большой полуоси. У германия она совпадает с направлением осей $[111]$ и на первую зону Бриллюэна приходится половина каждого эллипсоида энергии, а следовательно, в германии имеется не восемь, а только четыре полных эллипсоида энергии (рис. 2.22).

Абсолютный минимум зоны проводимости у кремния лежит в направлении осей $[100]$ недалеко от границы зоны Бриллюэна (рис. 2.21). Поэтому у кремния имеется шесть эквивалентных минимумов энергии, а следовательно, у кремния на первую зону Бриллюэна приходится шесть эллипсоидальных поверхностей постоянной энергии, вытянутых вдоль осей $[100]$ (рис. 2.22). Центры эллипсоидов расположены на расстоянии трех четвертей от центра зоны Бриллюэна.

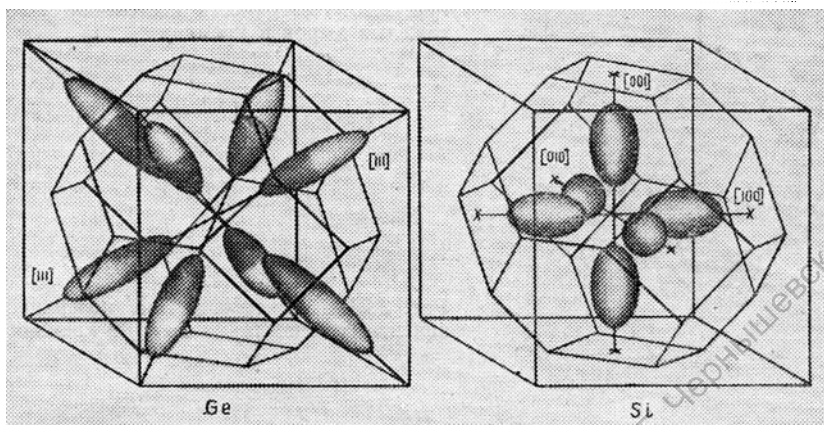


Рис. 2.22. Формы поверхностей постоянной энергии в зонах проводимости германия и кремния

У арсенида галлия (см. рис. 2.21) и антимонида индия (рис. 2.25) абсолютный минимум зоны проводимости лежит при $k = 0$, поэтому у них поверхностями постоянной энергии являются сферы с центром в центре зоны Бриллюэна и эффективные массы электронов соответственно равны $0,068m_0$ и $0,013m_0$.

Зависимость энергии электронов от волнового вектора k вблизи абсолютных минимумов зоны проводимости в германии и кремнии выражается формулой

$$E(k) = E(k_0) + \frac{\hbar^2 [(k_x - k_{0x})^2 + (k_y - k_{0y})^2]}{2m_1} + \frac{\hbar^2 (k_z - k_{0z})^2}{2m_2}. \quad (2.152)$$

Значения компонент тензора эффективной массы электрона, определенные по циклотронному резонансу, составляют у германия $m_1 = m_2 = m_3 = 0,082 m_0$; $m_3 = m_1 = 1,59 m_0$, а у кремния $m_1 = m_2 = m_3 = 0,19m_0$; $m_3 = m_1 = 0,92m_0$. Отношение m_3/m_1 характеризует анизотропию свойств изоэнергетических поверхностей. Для германия это отношение равно 19,3, а у кремния — 4,8.

Минимальное расстояние между дном зоны проводимости и дном валентной зоны называют шириной запрещенной зоны. У кремния и германия экстремумы энергии электронов и дырок лежат в различных точках зоны Бриллюэна (см. рис. 2.21). При комнатной температуре и нормальном атмосферном давлении ширина запрещенной зоны германия достаточной степени чистоты равна 0,67 эВ, кремния — 1,12 эВ, арсенида галлия — 1,43 эВ и антимонида индия — 0,18 эВ.

В приближении сильной связи валентная зона германия и кремния образуется из атомных p -орбиталей, а поэтому валентная зона их включает три подзоны, которые условно можно обозначить как $P_{(3/2)}^{(3/2)-}$, $P_{(3/2)}^{(1/2)-}$ и $P_{(1/2)}^{(1/2)-}$ зоны. Для всех трех зон максимум энергии на-

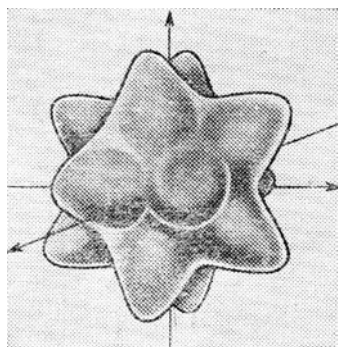


Рис. 2.23. Поверхность постоянной энергии тяжелых дырок в кремнии

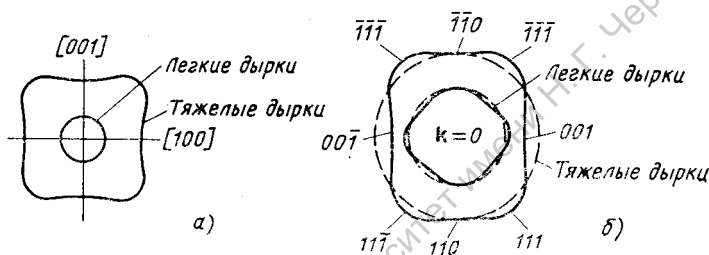


Рис. 2.24. Изоэнергетические поверхности в валентных зонах германия (а) и кремния (б)

ходится в центре зоны Бриллюэна $k = 0$ (см. рис. 2.21). При этом обе $P_{3/2}$ -зоны вырождены при $k = 0$. Для $P_{1/2}$ -зоны вырождение частично снимается за счет спин-орбитального взаимодействия, в результате этого она расположена ниже на величину E_{so} .

У обеих $P_{3/2}$ -зон при $k \neq 0$ вырождение снимается и расчет зависимости $E(k)$ для них приводит к выражению

$$E_{1,2}(k) = E(0) - \frac{\hbar^2}{2m_0} \left[Ak^2 \pm \sqrt{B^2 k^4 + C^2 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)} \right], \quad (2.153)$$

где A, B, C — безразмерные константы, равные соответственно для германия $13,1 \pm 0,4$; $8,3 \pm 0,6$; $12,5 \pm 0,5$ и для кремния $4,0 \pm 0,1$; $1,1 \pm 0,4$; $4,1 \pm 0,4$. Поверхности равной энергии (2.153) представляют собой гофрированные поверхности (рис. 2.23 и 2.24).

Усреднение зависимости (2.153) по различным направлениям в k -пространстве позволяет заменить гофрированную поверхность сферической (рис. 2.24), для которой $E_{1,2}$ может быть представлено выражением

$$E_{1,2} = E(0) - \frac{\hbar^2 k^2}{2m_0} \left[A \pm \sqrt{B^2 + \frac{1}{5} C^2} \right]. \quad (2.154)$$

В этом случае эффективная масса является скалярной величиной

$$m_p^* = \frac{m_0}{A \pm \sqrt{B^2 + C^2/5}} \quad (2.155)$$

и должны существовать два вида дырок:
тяжелые

$$m_{pT}^* = \frac{m_0}{A - \sqrt{B^2 + C^2/5}} \quad (2.156)$$

и легкие

$$m_{pL}^* = \frac{m_0}{A + \sqrt{B^2 + C^2/5}}. \quad (2.157)$$

Согласно рис. 2.21 зоной тяжелых дырок является широкая зона (с малым значением $\frac{\partial^2 E}{\partial k^2}$), а зоной легких дырок — более узкая (с большим значением $\frac{\partial^2 E}{\partial k^2}$).

Из экспериментов по циклотронному резонансу найдены следующие значения эффективных масс тяжелых и легких дырок:
для германия

$$m_{pT}^* = 0,33m_0; \quad m_{pL}^* = 0,04m_0;$$

для кремния

$$m_{pT}^* = 0,56m_0; \quad m_{pL}^* = 0,16m_0;$$

для арсенида галлия

$$m_{pT} = 0,50m_0; \quad m_{pL} = 0,12m_0.$$

Зависимость $E(k)$ для третьей ветви валентной зоны имеет вид:

$$E_3(k) = E(0) - E_{s0} - \frac{\hbar^2 k^2}{2m_0} A, \quad (2.158)$$

где E_{s0} — энергия спин-орбитального взаимодействия, равная для германия 0,3 и для кремния 0,04 эВ. Поверхности равной энергии третьей валентной зоны согласно (2.158) представляют собой сферы. Поэтому эффективная масса третьего типа дырок является скалярной величиной и равна $m_p^* = m_0/A$.

Как показывают эксперименты, зона проводимости антимонида индия обладает сферической симметрией и минимум ее находится в центре зоны Бриллюэна (рис. 2.25). Вблизи минимума энергия подчиняется параболическому закону, а кривизна зоны очень велика. В силу этого эффективная масса электронов, находящихся на дне этой зоны, очень мала и составляет 0,013 m_0 по данным измерений из циклотронного резонанса. Для больших энергий зона проводимости имеет непараболическую форму и кривизна ее уменьшается. Это подтверждают данные рис. 2.26, из которого следует, что энергия непропорциональна k^2 . В этом случае энергия вблизи минимума так называемой нестандартной зоны имеет вид:

$$E = E_c + \frac{\hbar^2 k^2}{2m_n^*} (1 - \alpha k^2), \quad (2.159)$$

где α — параметр, определяющий степень отклонения зоны проводимости от параболическости; E_c — энергия дна зоны проводимости; m_n^* — эффективная масса электрона.

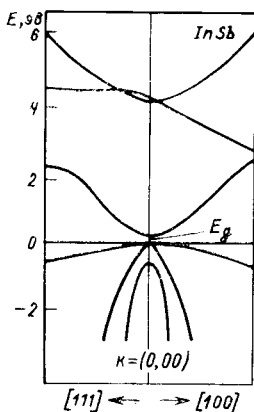


Рис. 2.25. Энергетическая зонная структура антимонида индия

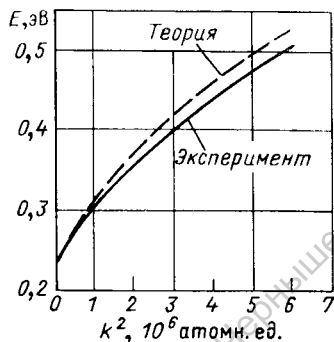


Рис. 2.26. Отклонение формы зоны проводимости в антимониде индия от параболической

Валентная зона антимонида индия, зонная структура которого изображена на рис. 2.25, включает три ветви $E(\mathbf{k})$. Зона тяжелых дырок V_1 ($m_{pT}^* = 0,6m_0$) и зона легких дырок V_2 ($m_{pL}^* = 0,012m_0$) вырождены при $\mathbf{k} = 0$. Третья зона отщеплена вследствие спин-орбитального взаимодействия на $E_{s0} = 0,9$ эВ.

2.14. МЕТОД ЭФФЕКТИВНОЙ МАССЫ

При легировании полупроводника донорными и акцепторными примесями в нем появляются локальные состояния, характеризующиеся примесными уровнями. Но прежде чем перейти к изложению теории локализованных состояний электрона в кристалле, рассмотрим один из методов приближенного решения уравнения Шредингера в случае, когда на периодический потенциал поля решетки $V(\mathbf{r})$ наложено некоторое внешнее возмущение $U(\mathbf{r})$. Уравнение Шредингера в этом случае имеет вид:

$$\left[-\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) + U(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2.160)$$

Прямое решение этого уравнения затруднено тем, что нам неизвестен потенциал $V(\mathbf{r})$.

Из предыдущего рассмотрения следует, что уравнение движения электрона в периодическом поле решетки при наличии внешнего электрического поля аналогично классическому уравнению движения свободного электрона, если его массу m_0 заменить в общем случае тензором эффективной массы m^* .

Уравнение Шредингера для свободного электрона имеет вид:

$$-\frac{\hbar^2}{2m_0} \Delta \Psi_0(\mathbf{r}) = E \Psi_0(\mathbf{r}). \quad (2.161)$$

Тогда для электрона в кристалле уравнение Шредингера можно записать аналогично (2.161):

$$-\frac{\hbar^2}{2m^*} \Delta \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2.162)$$

Необходимо выяснить, при каких условиях (2.162) дает решения, подобные решениям уравнения Шредингера с периодическим потенциалом $V(\mathbf{r})$:

$$\left[-\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2.163)$$

Известно, что уравнению (2.163) удовлетворяют блоховские функции $\Psi(\mathbf{r}) = U_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}}$, а энергетический спектр электрона вблизи экстремальной точки ($\mathbf{k} = \mathbf{k}_0$) описывается выражением

$$E = E(\mathbf{k}_0) + \frac{\hbar^2}{2} \left[\frac{(k_x - k_{0x})^2}{m_1} + \frac{(k_y - k_{0y})^2}{m_2} + \frac{(k_z - k_{0z})^2}{m_3} \right]. \quad (2.164)$$

В уравнении (2.164) предполагается, что тензор обратной эффективной массы приведен к диагональному виду, и поэтому (2.162) может быть записано в виде

$$\sum_{i=1}^3 \frac{1}{m_i} \frac{\partial^2 \Psi}{\partial x_i^2} + \frac{2E}{\hbar^2} \Psi = 0. \quad (2.165)$$

Нетрудно убедиться, что решением этого уравнения будет:

$$\Psi = A e^{i(\alpha x + \beta y + \gamma z)}, \quad (2.166)$$

при этом

$$E = \frac{\hbar^2}{2} \left[\frac{\alpha^2}{m_1} + \frac{\beta^2}{m_2} + \frac{\gamma^2}{m_3} \right]. \quad (2.167)$$

Сравнивая выражения (2.164) и (2.167), можно сделать вывод, что если положить $\alpha = k_x - k_{0x}$; $\beta = k_y - k_{0y}$; $\gamma = k_z - k_{0z}$, а начало отсчета энергии вести от экстремальной точки, т. е. $E(\mathbf{k}_0) = 0$, то решение уравнения (2.162) дает тот же спектр энергии электрона в кристалле вблизи экстремальной точки $\mathbf{k} = \mathbf{k}_0$, что и уравнение (2.163).

Таким образом, для решения задачи определения спектра энергии электрона кристалла при наличии внешних полей можно вместо (2.160) решать уравнение

$$\left[-\frac{\hbar^2}{2m^*} \Delta + U(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}), \quad (2.168)$$

в которое уже не входит неизвестный потенциал поля решетки $V(\mathbf{r})$. Подобный метод решения получил название метода эффективной массы. Большое достоинство этого метода состоит в том, что значения компонент тензора эффективной массы вблизи экстремальной точки можно определить из эксперимента (например, из опытов по циклотронному резонансу). То, что метод эффективной массы справедлив лишь для электронов вблизи экстремума энергии, не накладывает серьезных ограничений при решении большинства задач теории полупроводников, поскольку концентрация

электронов (или дырок) в невырожденных полупроводниках много меньше числа состояний в зоне и, следовательно, электроны (или дырки) располагаются в разрешенных зонах в состояниях с минимальной энергией.

2.15. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ПРИМЕСНЫХ СОСТОЯНИЙ

До сих пор рассматривались идеальные кристаллы, обладающие периодической структурой. Теперь рассмотрим полупроводники, в решетке которых имеются простейшие дефекты в виде примесных атомов V или III группы таблицы Менделеева.

Предположим, что один из узлов в кристалле кремния замещен атомом мышьяка, имеющим пять валентных электронов. Четыре из них будут участвовать в образовании ковалентной связи с ближайшими соседними атомами кремния. Пятый электрон, не принимая участия в образовании ковалентной связи, будет взаимодействовать с большим числом атомов кремния и в результате окажется сравнительно слабо связанным с ионом мышьяка силами кулоновского взаимодействия. Его поведение подобно состоянию электрона в атоме водорода. Однако при решении задачи о потенциале ионизации электрона примеси следует учесть дополнительно два обстоятельства. Во-первых, этот электрон находится не только в кулоновском поле иона мышьяка, но и в периодическом поле кристаллической решетки. В связи с этим при описании его движения необходимо пользоваться не действительной, а эффективной массой m^* . Во-вторых, взаимодействие между электроном и положительным ионом мышьяка с зарядом Ze происходит в твердом теле, обладающем относительной диэлектрической проницаемостью ϵ_r , поэтому в выражение для потенциальной энергии $U(\mathbf{r})$ электрона примеси следует ввести ϵ_r , тогда

$$U(\mathbf{r}) = -\frac{Ze^2}{4\pi\epsilon_r\epsilon_0 r}, \quad (2.169)$$

где ϵ_0 — электрическая постоянная.

Соотношение (2.169) справедливо только для макроскопических точечных зарядов, т. е. оно справедливо в том случае, если расстояние между взаимодействующими зарядами достаточно велико по сравнению с постоянной решетки кристалла, т. е. орбита электрона примеси должна охватывать большое число узлов кристаллической решетки.

С учетом сказанного уравнение Шредингера для пятого электрона атома мышьяка, не участвующего в образовании ковалентной связи, согласно (2.161) примет вид:

$$\left(-\frac{\hbar^2}{2m^*}\Delta - \frac{Ze^2}{4\pi\epsilon_r\epsilon_0 r}\right)\Psi_a = E_n\Psi_a. \quad (2.170)$$

Решим это уравнение по аналогии с уравнением для атома водорода, получим выражение для собственных значений энергии электрона донорной примеси:

$$E_n = E_c - \frac{m^*Z^2e^4}{8\hbar^2\epsilon_r^2\epsilon_0^2} \frac{1}{n^2}, \quad (2.171)$$

причем энергия электрона отсчитывается от дна зоны проводимости, а n — квантовое число, определяющее различные уровни возбуждения донорной примеси ($n \geq 1$).

Эту формулу можно записать следующим образом:

$$E_n = E_c - \frac{m_0 Z^2 e^4}{8 \hbar^2 \epsilon_0^2} \left(\frac{m^*}{m_0} \right) \frac{1}{\epsilon_r^2 n^2}. \quad (2.172)$$

Если подставить численные значения e , m_0 , $\hbar$, ϵ_0 и выразить энергию в электрон-вольтах, то получим:

$$E_n = E_c - \frac{13,52 Z^2}{\epsilon_r^2} \left(\frac{m^*}{m_0} \right) \frac{1}{n^2} = E_c - \frac{E_d}{n^2}, \quad (2.173)$$

где 13,52 соответствует значению энергии ионизации атома водорода в электрон-вольтах, а

$$E_d = \frac{13,52 Z^2}{\epsilon_r^2} \left(\frac{m^*}{m_0} \right) \quad (2.174)$$

есть энергия основного состояния ($n = 1$) атома донорной примеси.

Из (2.173) следует, что *уровень энергии донорной примеси, соответствующий основному состоянию ($n = 1$), лежит в запрещенной зоне полупроводника ниже дна зоны проводимости E_c на величину E_d .*

Энергия ионизации донорной примеси, равная E_d , согласно равенству (2.174) в ϵ_r^2 раз меньше энергии ионизации атома водорода. Кроме того, она зависит от Z^2 . Это означает, что уровень двукратно заряженного иона донорной примеси должен быть расположен в запрещенной зоне ниже уровня однократно заряженного иона этой примеси.

Если положить $m^* = 0,25 m_0$, то в случае германия, у которого $\epsilon_r = 16$, энергия ионизации пентавалентных атомов примеси должна быть примерно 0,01 эВ. Значения энергии ионизации донорных примесей, полученные экспериментально, достаточно хорошо согласуются с рассчитанным значением (рис. 2.27).

Расчет радиуса орбиты электрона примеси по водородоподобной модели дает формулу

$$a = \frac{\epsilon_0 \hbar^2}{\pi e^2 m^*} \epsilon_r n^2 = a_0 \left(\frac{m_0}{m^*} \right) \epsilon_r n^2, \quad (2.175)$$

где $a_0 = \epsilon_0 \hbar^2 / \pi e^2 m_0 = 0,053$ нм — радиус первой борвской орбиты атома водорода.

Радиус первой орбиты электрона донорной примеси в германии составляет: $a_1 = 64 a_0 = 3,4$ нм. Если учесть, что постоянная решетки германия равна 0,356 нм, то эта орбита охватывает приблизительно 200 узлов решетки. Это подтверждает возможность рассмотрения движения электрона примеси в среде с постоянной ϵ_r .

Рассмотрим теперь случай акцепторной примеси. Допустим, что один из узлов решетки кремния замещен атомом бора — элементом третьей группы периодической системы. В этом случае у примесного атома не хватает одного электрона для заполнения ковалентной

Элементы I, II, VI и VIII групп, введенные в полупроводник в качестве легирующих примесей, создают «глубокие» уровни энергии. Эти элементы могут входить в решетку в виде атома (донора D или акцептора A), однократно заряженного иона [$D^{(1)}$ или $A^{(1)}$] и двукратно заряженного иона [$D^{(2)}$ или $A^{(2)}$]. Кроме того, одна и та же примесь может создавать как донорные, так и акцепторные уровни.

На рис. 2.27 показаны измеренные значения энергии ионизации для различных примесей в германии и кремнии. Примесные уровни, лежащие выше середины запрещенной зоны, являются донорными. Донорные уровни, лежащие ниже середины запрещенной зоны, отмечены буквой D . Энергия ионизации примесей, уровни которых лежат ниже середины запрещенной зоны, отсчитывается от потолка валентной зоны. Такие уровни являются акцепторными. Акцепторные уровни, лежащие выше середины запрещенной зоны, отмечены буквой A .

ГЛАВА ТРЕТЬЯ

КОЛЕБАНИЯ АТОМОВ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ

3.1. ОДНОМЕРНЫЕ КОЛЕБАНИЯ ОДНОРОДНОЙ СТРУНЫ

Как известно, атомы в кристалле не неподвижны, они совершают тепловые колебания относительно своих положений равновесия. Эти колебания нарушают строгую периодичность поля решетки. Рассеяние носителей заряда на этих нерегулярностях решетки оказывает существенное влияние на многие физические явления в полупроводниках (см. гл. 5 и 6).

Теория колебаний атомов трехмерного кристалла достаточно сложна. Вначале рассмотрим колебания атомов одномерной решетки, а затем полученные закономерности распространим на объемный кристалл.

Но прежде рассмотрим распространение продольных волн в одномерной однородной упругой струне, линейная плотность которой ρ . Допустим, что на участок струны Δx действует сила. Предположим, что вызванное этой силой продольное смещение точки с координатой x равно $u(x)$. Тогда деформация, определяемая как относительное изменение длины на конце элемента с координатой x , будет:

$$e(x) = \frac{\partial u}{\partial x}, \quad (3.1)$$

а на другом конце участка струны с координатой $x + \Delta x$

$$e(x + \Delta x) = e(x) + \frac{\partial^2 u}{\partial x^2} \Delta x. \quad (3.2)$$

Если сила, вызывающая деформацию, равна F , то модуль упругости T_n определяется отношением $T_n = F/e$. Поэтому результи-

рующая сила $F_{\Delta x}$, действующая на элемент струны Δx , равна:

$$F_{\Delta x} = T_n [e(x + \Delta x) - e(x)] = T_n \frac{\partial^2 u}{\partial x^2} \Delta x. \quad (3.3)$$

Эту силу можно определить как произведение массы элемента струны $\rho \Delta x$ на ускорение $\partial^2 u / \partial t^2$, т. е.

$$F_{\Delta x} = \rho \frac{\partial^2 u}{\partial t^2} \Delta x. \quad (3.4)$$

В результате получим уравнение свободных колебаний струны:

$$\frac{\partial^2 u}{\partial t^2} = \frac{T_n}{\rho} \frac{\partial^2 u}{\partial x^2}. \quad (3.5)$$

Здесь

$$\sqrt{T_n / \rho} = v_{зв} \quad (3.6)$$

есть скорость распространения упругой волны (звуковой волны), которая не зависит от частоты колебания.

Решение уравнения колебания струны имеет вид бегущей монохроматической волны

$$u = A e^{i(qx - \omega t)}, \quad (3.7)$$

где A — амплитуда колебания; ν — частота колебания; $\omega = 2\pi\nu$ — круговая (циклическая, угловая) частота колебания; λ — длина волны; $q = 2\pi/\lambda$ — волновое число.

Для упругой волны, распространяющейся в непрерывной струне, частота колебаний

$$\omega = 2\pi\nu = 2\pi v_{зв} / \lambda = v_{зв} q \quad (3.8)$$

пропорциональна волновому числу q .

При этом абсолютное значение волнового числа q может изменяться от 0 до ∞ , а следовательно, частота колебаний изменяется непрерывно от нуля до ∞ (см. рис. 3.2). Каждому значению q соответствует волна, характеризующаяся определенной частотой колебаний.

3.2. КОЛЕБАНИЯ ОДНОАТОМНОЙ ЛИНЕЙНОЙ ЦЕПОЧКИ

Рассмотрим теперь линейную цепочку одинаковых атомов массы m , расположенных в состоянии равновесия на расстоянии a друг от друга (рис. 3.1). Пусть n -й атом смещен на расстояние u_n . Если отклонение этого атома от положения равновесия мало, то силу взаимодействия можно рассматривать как квазиупругую, т. е. пропорциональную изменению расстояния между атомами.

Поэтому при учете лишь ближайших соседей сила F будет равна сумме сил, действующих со стороны атомов $n-1$ и $n+1$:

$$F_n = -\beta(u_n - u_{n+1}) - \beta(u_n - u_{n-1}) = \beta(u_{n+1} + u_{n-1} - 2u_n), \quad (3.9)$$

где β — коэффициент квазиупругой силы.

Уравнение движения n -го атома имеет вид:

$$m \frac{d^2 u_n}{dt^2} = \beta (u_{n+1} + u_{n-1} - 2u_n). \quad (3.10)$$

Решение уравнения будем искать в виде

$$u_n = A e^{i(qan - \omega t)}. \quad (3.11)$$

Подставляя (3.11) в (3.10), получаем:

$$-m\omega^2 = \beta (e^{iqa} + e^{-iqa} - 2). \quad (3.12)$$

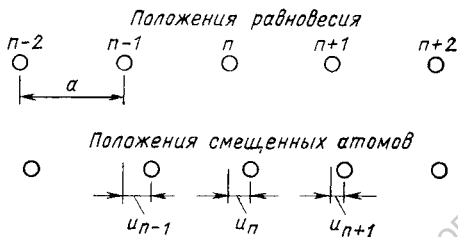


Рис. 3.1. Схематическое изображение линейной решетки из одинаковых атомов

Но так как $2 - e^{iqa} - e^{-iqa} = 2(1 - \cos qa) = 4 \sin^2 \frac{aq}{2}$, то

$$\omega = \pm 2(\beta/m)^{1/2} \sin \frac{aq}{2} = \pm \omega_m \sin \frac{aq}{2}, \quad (3.13)$$

где

$$\omega_m = 2(\beta/m)^{1/2}. \quad (3.14)$$

Зависимость частоты ω от волнового числа q носит название закона дисперсии для колебаний. Таким образом, решения (3.11) типа упругой волны удовлетворяют уравнениям движения (3.10) для любого n , если частота колебаний волны связана с волновым числом q дисперсионным соотношением вида (3.13).

Скорость распространения упругой волны вдоль линейной цепочки атомов

$$v = \frac{\omega \lambda}{2\pi} = \frac{\lambda}{\pi} (\beta/m)^{1/2} \sin \frac{\pi a}{\lambda} \quad (3.15)$$

зависит от длины волны. Это является специфическим свойством упругих волн в среде с дискретной структурой.

Определим пределы изменения волнового числа q . Для этого сравним значения смещения (3.11) для волновых чисел q и $q' = q + \frac{2\pi}{a}g$, где g — целое положительное или отрицательное число.

В этом случае

$$(u_n)_g = A e^{i(qan - \omega t)} e^{i2\pi ng} = A e^{i(qan - \omega t)} = u_n, \quad (3.16)$$

так как n — порядковый номер атома, а ng — целое число и $e^{i2\pi ng} = 1$.

Из соотношения (3.16) следует, что волна $(u_n)_g$ во всех точках и во все моменты времени совпадает с волной u_n . Это значит, что волновые числа q и q' физически неразличимы и, следовательно, характеризуют одно и то же колебание. Поэтому достаточно рассматривать дисперсионное соотношение при изменении волнового числа в пределах

$$-\frac{\pi}{a} \leq q \leq \frac{\pi}{a}. \quad (3.17)$$

Отсюда можно сделать вывод, что область, в которой лежат волновые векторы $\mathbf{q}$, совпадает с зоной Бриллюэна для волнового вектора $\mathbf{k}$ электронов. Поэтому значения $\mathbf{q}$ можно рассматривать в том же пространстве, что и $\mathbf{k}$ (или $\mathbf{p}$).

Для определения значения q поступим следующим образом. Расположим атомы в одномерной цепочке по окружности так, чтобы N -й и 1-й атомы находились в равновесии на расстоянии постоянной решетки a . Так как число атомов N велико, то радиус окружности будет большим. Поскольку в цепочке взаимодействуют лишь атомы, находящиеся на расстоянии нескольких постоянных решетки, то условия, в которых будут находиться граничные атомы, не влияют на движение атомов внутри цепочки. Для этого случая граничные условия можно заменить условиями цикличности Борна—Кармана, согласно которым

$$u_{n \pm N} = u_n, \quad (3.18)$$

так как в этом случае атом номера $n \pm N$ совпадает с номером атома n .

Подставив (3.18) в (3.11), получим:

$$u_{n \pm N} = A e^{i[qa(n \pm N) - \omega t]} = A e^{(qaN - \omega t)} e^{\pm iqaN} = u_n. \quad (3.19)$$

Из (3.19) следует, что $e^{\pm iqaN} = 1$. Это значит, что

$$qaN = 2\pi g, \quad (3.20)$$

откуда

$$q = \frac{2\pi}{aN} g. \quad (3.21)$$

С учетом (3.17) найдем, что g должно меняться в пределах

$$-\frac{N}{2} \leq g \leq \frac{N}{2}. \quad (3.22)$$

Таким образом, для атомной цепочки, имеющей N одинаковых атомов, т. е. N степеней свободы, волновой вектор $\mathbf{q}$ меняется дискретно и в интервале от $-\pi/a$ до π/a может принимать N значений, т. е. столько значений, сколько элементарных ячеек содержится в цепочке. Поскольку N велико, то изменения $\mathbf{q}$ можно рассматривать как квазинепрерывные.

Для линейной решетки, состоящей из одинаковых атомов, закон дисперсии представлен на рис. 3.2. На этом же рисунке дана зависимость $\omega(\mathbf{q})$ для упругой однородной струны (континуум).

Для длинных волн, т. е. малых значений q , можно положить $\sin \frac{aq}{2} \approx \frac{aq}{2}$. Тогда из (3.13) будем иметь:

$$\omega = 2(\beta/m)^{1/2} \frac{aq}{2} = \sqrt{\beta/m} a q. \quad (3.23)$$

Для линейной атомной цепочки линейная плотность $\rho = m/a$, а модуль упругости T_n , равный отношению силы к относительному изменению смещения

$$T_n = \frac{F_n}{|u_n - u_{n+1}|/a} = \beta a, \quad (3.24)$$

поэтому

$$v_{зв} = \sqrt{T_{н}/\rho} = \sqrt{\beta/m} a. \quad (3.25)$$

С учетом (3.25) соотношение (3.23) принимает вид:

$$\omega = v_{зв} q, \quad (3.26)$$

а максимальная частота ω_m равна:

$$\omega_m = 2(\beta/m)^{1/2} = 2v_{зв}/a. \quad (3.27)$$

Таким образом, для длинных волн частота колебаний одноатомной цепочки пропорциональна волновому числу q .

В случае дисперсии волны фазовая скорость, с которой распространяется фаза монохроматической волны, $v_{\phi} = \omega/q$ и групповая скорость, с которой распространяется волновой пакет, а следовательно, энергия волны, $v_{гр} = d\omega/dq$ (рис. 3.3). Для линейной цепочки одинаковых атомов согласно (3.13) и (3.27)

$$v_{\phi} = \omega_m \left| \frac{\sin \frac{aq}{2}}{q} \right| = v_{зв} \left| \frac{\sin \frac{aq}{2}}{\frac{1}{2} aq} \right|; \quad (3.28)$$

$$v_{гр} = \frac{a\omega_m}{2} \left| \cos \frac{aq}{2} \right| = v_{зв} \left| \cos \frac{aq}{2} \right|. \quad (3.29)$$

Для длинных волн, т. е. малых значений q ,

$$v_{\phi} = v_{гр} = v_{зв}, \quad (3.30)$$

фазовая и групповая скорости равны и представляют собой скорость звука. Это означает, что в пределах длинных волн одноатомная цепочка ведет себя как упругая струна.

Как следует из рис. 3.2, с увеличением значений q , т. е. с уменьшением длины волны, частота колебаний одноатомной цепочки отклоняется от линейной зависимости. Минимальная длина волны соответствует максимальному значению волнового числа $q_m = \pi/a$ и $\lambda_{min} = 2a$. Физически это положение представляется ясным, так

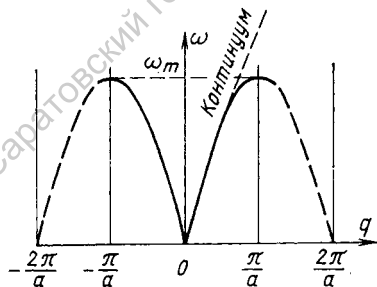


Рис. 3.2. Зависимость частоты колебания от волнового числа для линейной решетки, состоящей из одинаковых атомов, и для упругой струны

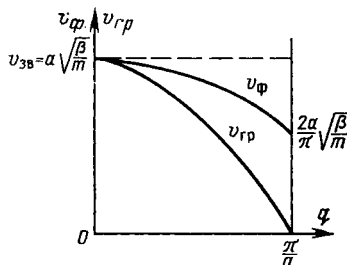


Рис. 3.3. Зависимость фазовой и групповой скорости от волнового числа

как в дискретной цепочке не могут существовать волны с длиной полуволны, меньшей, чем период решетки a . Волне с наименьшей длиной $\lambda_{min} = 2a$ соответствует максимальная частота ω_m . Таким образом, для колебаний дискретных атомных структур характерным является существование максимальной частоты. В твердых телах скорость звука достигает $3 \cdot 10^5$ см/с, а постоянная решетки составляет примерно $3 \cdot 10^{-8}$ см, поэтому наибольшая частота колебаний, возможная в решетке, $\omega_m = 2 \cdot 3 \cdot 10^5 / 3 \cdot 10^{-8} = 2 \cdot 10^{13}$ Гц и лежит в инфракрасной области.

Для самых коротких длин волн, т. е. для $q_m = \pi/a$, фазовая скорость $v_\phi = 2v_{зв}/\pi = \frac{2a}{\pi} \sqrt{\beta/m}$, а групповая скорость, с которой переносится энергия колебаний решетки, уменьшается до нуля. Это свидетельствует о том, что эти моды колебаний характеризуют стоячие волны в решетке.

3.3. ЭНЕРГИЯ КОЛЕБАНИЙ АТОМОВ ОДНОМЕРНОЙ РЕШЕТКИ. НОРМАЛЬНЫЕ КООРДИНАТЫ

В атомной цепочке выражение (3.11) удовлетворяет уравнению движения (3.10), если выполнено соотношение дисперсии (3.13). Но вследствие действия мощных сил связи колебание какого-либо атома передается соседним атомам и в кристалле фактически возбуждается коллективное движение в форме волны, охватывающее все атомы кристалла. Поэтому в случае малых колебаний движение атомов решетки может быть представлено посредством линейной суперпозиции (суммы) бегущих монохроматических волн смещения (3.11). Каждая из этих волн будет характеризоваться волновым числом q , соответствующей ему частотой ω_q и амплитудой A_q . В этом случае смещение n -го атома цепочки запишется в виде

$$u_n = \sum_q A_q e^{i(qan - \omega_q t)}. \quad (3.31)$$

Введем новые координаты, положив

$$Q_q = C_q e^{-i\omega_q t}, \quad (3.32)$$

где $C_q = \sqrt{N} A_q$.

Независимые переменные Q_q , изменяющиеся во времени по гармоническому закону, называют нормальными, или главными, координатами решетки кристалла.

Подставим значения Q_q и A_q в (3.31), в результате будем иметь:

$$u_n = \frac{1}{\sqrt{N}} \sum_q Q_q e^{i q a n}. \quad (3.33)$$

Определим теперь внутреннюю (полную) энергию колебаний атомов линейной решетки. Она складывается из кинетической T и потенциальной U энергии всех атомов. Кинетическая энергия согласно (3.33) равна:

$$T = \sum_{n=1}^N \frac{m}{2} \dot{u}_n^2 = \frac{m}{2N} \sum_q \dot{Q}_n^2 \sum_{n=1}^N e^{i2qan}. \quad (3.34)$$

Подсчитаем сомножитель $\sum_{n=1}^N e^{i2qan}$, учитывая выражение (3.21). Он вычисляется как сумма геометрической прогрессии

$$\begin{aligned} \sum_{n=1}^N e^{i2qan} &= \sum_{n=1}^N e^{i \frac{4\pi}{N} ng} = e^{i \frac{4\pi}{N} g \cdot 1} + e^{i \frac{4\pi}{N} g \cdot 2} + \dots + e^{i \frac{4\pi}{N} g N} = \\ &= \frac{e^{i \frac{4\pi}{N} g} - e^{i \frac{4\pi}{N} g} e^{i4\pi g}}{1 - e^{i \frac{4\pi}{N} g}} = \frac{e^{i \frac{4\pi}{N} g} (1 - e^{i4\pi g})}{1 - e^{i \frac{4\pi}{N} g}} = \\ &= \begin{cases} 0, & \text{когда } q \neq 0 \text{ (или } g \neq 0); \\ N, & \text{когда } q = 0 \text{ (или } g = 0), \end{cases} \quad (3.35) \end{aligned}$$

так как $e^{i4\pi g} = 1$, если $g \neq 0$; а сумма ряда равна N , если $g = 0$.

С учетом (3.35) кинетическая энергия атомов одномерной решетки примет вид:

$$T = \sum_q \frac{m}{2} \dot{Q}_q^2. \quad (3.36)$$

Так как сила $F_n = -\frac{\partial U}{\partial u_n}$, то для потенциальной энергии решетки U с учетом (3.9) и (3.11) получим:

$$U = U_0 + \frac{\beta}{2} \sum_{n=1}^N (2u_n - u_{n+1} - u_{n-1}) u_n. \quad (3.37)$$

Для определения β используем уравнение движения (3.10), в которое подставим смещение (3.31):

$$m \sum_q \omega_q^2 A_q e^{i(qan - \omega_q t)} = \beta \sum_q (2 - e^{iqa} - e^{-iqa}) A_q e^{i(qan - \omega_q t)}.$$

Отсюда найдем:

$$\beta = \sum_q \frac{m \omega_q^2}{2 - e^{iqa} - e^{-iqa}}. \quad (3.38)$$

Теперь в (3.37) подставим значения β (3.38) и смещения (3.33), выраженные через нормальные координаты. В результате, используя (3.35), будем иметь:

$$U = U_0 + \sum_q \frac{m}{2} \omega_q^2 Q_q^2. \quad (3.39)$$

Полная энергия колебаний решетки

$$E = T + U = U_0 + \sum_q \left(\frac{m}{2} \dot{Q}_q^2 + \frac{m}{2} \omega_q^2 Q_q^2 \right). \quad (3.40)$$

При $U_0 = 0$

$$E = \sum_q \left(\frac{m}{2} \dot{Q}_q^2 + \frac{m}{2} \omega_q^2 Q_q^2 \right) = \sum_q E_q. \quad (3.41)$$

Выражение

$$E_q = \frac{m}{2} \dot{Q}_q^2 + \frac{m}{2} \omega_q^2 Q_q^2 \quad (3.42)$$

совпадает с выражением полной энергии гармонического осциллятора с массой m . Спектр разрешенных значений энергии E_q осциллятора согласно квантовой механике определяется равенством

$$E_q = \hbar \omega_q \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (3.43)$$

Следовательно, энергия решетки может быть представлена в виде суммы энергии независимых гармонических осцилляторов

$$E = \sum_q E_q = \sum_q \hbar \omega_q \left(n + \frac{1}{2} \right). \quad (3.44)$$

Из (3.41) следует, что полная энергия колебаний решетки, выраженная через нормальные координаты, имеет вид энергии совокупности гармонических осцилляторов. Это означает, что одномерный кристалл, содержащий N атомов, можно рассматривать как динамическую систему с N степенями свободы, и ее движение можно описать с помощью независимых друг от друга N нормальных координат. Каждая нормальная координата описывает определенную конфигурацию всех атомов кристалла, совершающую колебания по простому гармоническому закону (3.32). Такое коллективное колебание атомов кристалла носит название нормального колебания решетки. При этом каждый атом кристалла участвует во всех N нормальных колебаниях решетки. Координата Q удовлетворяет уравнению

$$\ddot{Q}_q + \omega_q^2 Q_q = 0 \quad (q = 1, 2, \dots, N), \quad (3.45)$$

которое является уравнением движения гармонического осциллятора.

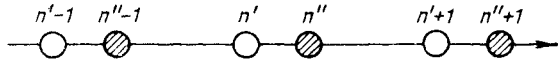
Таким образом, полная энергия теплового движения атомов одномерного кристалла складывается из энергий нормальных колебаний, ведущих себя подобно линейным гармоническим осцилляторам с собственной частотой ω_q .

3.4. КОЛЕБАНИЯ ДВУХАТОМНОЙ ЛИНЕЙНОЙ ЦЕПОЧКИ

Рассмотрим теперь колебания атомов кристалла, в элементарной ячейке которого имеется два атома. В качестве примеров таких веществ можно привести ионные щелочно-галогидные соединения, ковалентные полупроводники — германий, кремний, а также полупроводниковые соединения типа $A^{III}B^V$.

Предположим, что в одномерном кристалле, состоящем из атомов двух сортов с массами m_1 и m_2 (рис. 3.4), атомы m_1 находятся в узлах решетки n' , а атомы m_2 — в узлах n'' . Обозначим смещение n' -го

Рис. 3.4. Одномерная модель решетки, имеющей два атома в элементарной ячейке



и n'' -го атомов через u'_n и u''_n и будем считать, что коэффициент квазиупругой силы между соседними атомами одинаков и равен β .

Уравнения движения атомов обоих сортов имеют вид:

$$\left. \begin{aligned} m_1 \ddot{u}'_n &= \beta (u''_n + u''_{n-1} - 2u'_n); \\ m_2 \ddot{u}''_n &= \beta (u'_n + u'_{n+1} - 2u''_n). \end{aligned} \right\} \quad (3.46)$$

Решение этих уравнений будем искать в виде

$$\left. \begin{aligned} u'_n &= A_1 e^{i(qan - \omega t)}; \\ u''_n &= A_2 e^{i(qan - \omega t)}. \end{aligned} \right\} \quad (3.47)$$

Подставляя (3.47) в (3.46), получаем:

$$\left. \begin{aligned} (2\beta - m_1 \omega^2) A_1 &= \beta (1 + e^{-iqa}) A_2; \\ (2\beta - m_2 \omega^2) A_2 &= \beta (1 + e^{iqa}) A_1. \end{aligned} \right\} \quad (3.48)$$

Эта система уравнения относительно A_1 и A_2 имеет нетривиальное решение, если детерминант системы равен нулю. Он запишется в виде

$$\left\| \begin{array}{cc} 2\beta - m_1 \omega^2 & \beta(1 + e^{-iqa}) \\ \beta(1 + e^{iqa}) & 2\beta - m_2 \omega^2 \end{array} \right\| = 0. \quad (3.49)$$

Воспользовавшись равенствами $e^{iqa} + e^{-iqa} = 2 \cos aq$ и $1 - \cos aq = 2 \sin^2 \frac{aq}{2}$, будем иметь:

$$\omega^4 - 2\beta \frac{m_1 + m_2}{m_1 m_2} \omega^2 + \frac{4\beta^2}{m_1 m_2} \sin^2 \frac{aq}{2} = 0. \quad (3.50)$$

Из этого уравнения найдем:

$$\omega_1^2 = \frac{\omega_0^2}{2} \left[1 + \sqrt{1 - \gamma^2 \sin^2 \frac{aq}{2}} \right]; \quad (3.51)$$

$$\omega_2^2 = \frac{\omega_0^2}{2} \left[1 - \sqrt{1 - \gamma^2 \sin^2 \frac{aq}{2}} \right], \quad (3.52)$$

где

$$\omega_0^2 = 2\beta \frac{m_1 + m_2}{m_1 m_2}; \quad \gamma^2 = 4 \frac{m_1 m_2}{(m_1 + m_2)^2}. \quad (3.53)$$

Из формул (3.51) и (3.52) следует, что каждому значению q (или λ) соответствуют два значения ω (так как отрицательные значения ω не имеют смысла). Следовательно, у одномерного кристалла, в элементарной ячейке которого находится два атома, имеются две ветви колебаний; одну из них, характеризующуюся ω_1 , называют оптической, другую с ω_2 — акустической ветвью колебаний.

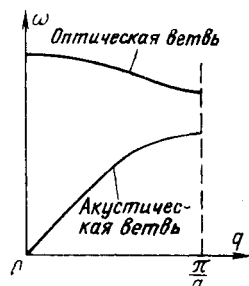
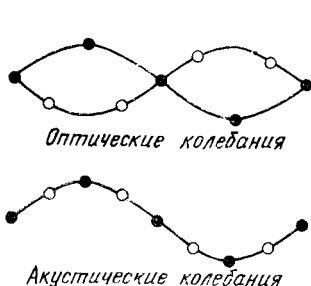


Рис. 3.5. Схематическое изображение колебаний частиц в двумерной линейной решетке, иллюстрирующее природу различия между оптическими и акустическими колебаниями. Предполагается, что колебания поперечные и что оба типа колебаний имеют одинаковую длину волны

Рис. 3.6. Оптические и акустические ветви колебаний одномерной решетки, в элементарной ячейке которой находятся два атома (положительные ветви кривых)

Рассмотрим два предельных случая: самые длинные и самые короткие волны. Для $q = 0$ и $q = \pi/a$ имеем:

$$\left. \begin{aligned} \omega_{\text{оп}}(0) &= \omega_0; & \omega_{\text{оп}}(\pi/a) &= \frac{\omega_0}{\sqrt{2}} \sqrt{1 + \sqrt{1 - \gamma^2}}; \\ \omega_{\text{ак}}(0) &= 0; & \omega_{\text{ак}}(\pi/a) &= \frac{\omega_0}{\sqrt{2}} \sqrt{1 - \sqrt{1 - \gamma^2}}. \end{aligned} \right\} \quad (3.54)$$

Отсюда следует, что

$$\omega_{\text{оп}}(0) = \omega_0 > \omega_{\text{оп}}(\pi/a) > \omega_{\text{ак}}(\pi/a) > \omega_{\text{ак}}(0) = 0. \quad (3.55)$$

Для бесконечно длинной волны акустической ветви колебаний, когда $\omega_{\text{ак}}(0) = 0$, согласно (3.47) и (3.48)

$$(u'_n/u''_n)_{\text{ак}} = (A_1/A_2)_{\text{ак}} = 1, \quad (3.56)$$

т. е. атомы смещаются одинаково.

Это означает, что в бесконечно длинной волне акустической ветви колебаний атомы движутся синхронно и в любой момент времени их отклонения от положения равновесия одинаковы, т. е. ячейки смещаются как одно целое (рис. 3.5).

Аналогичное положение имеет место при распространении упругих волн. Поэтому эта ветвь колебаний и была названа акустической.

Бесконечно длинной волне оптической ветви колебаний соответствует максимальная частота, равная ω_0 . Для нее согласно (3.47) и (3.48) справедливо отношение

$$(u'_n/u''_n)_{\text{оп}} = (A_1/A_2)_{\text{оп}} = -m_2/m_1. \quad (3.57)$$

Следовательно, в бесконечно длинной волне оптической ветви колебаний атомы ячейки смещаются в противоположных направлениях, они колеблются в противофазе, так что центр тяжести каждой пары атомов остается неподвижным, т. е. $m_1 u'_n + m_2 u''_n = 0$

(рис. 3.5). Если бы элементарную ячейку составляли ионы разного знака, то их колебания в противофазе вызвали бы смещение ионов и привели бы к изменению электрического дипольного момента ячейки. Это проявилось бы в поглощении и испускании инфракрасного излучения. Поэтому колебания этого типа и были названы оптическими.

Для длинных волн, а следовательно, для малых значений q , полагая $\sin(aq/2) \approx aq/2$, корни в (3.51) и (3.52) можно разложить в ряд. В результате для длинных волн оптической ветви колебаний

$$\omega_{\text{оп}} = \omega_0 \left(1 - \frac{\gamma^2 a^2}{32} q^2 \right). \quad (3.58)$$

При $q \rightarrow 0$ фазовая скорость оптических колебаний $v_{\text{ф}} = \frac{\omega_{\text{оп}}}{q}$ стремится к ∞ , а групповая скорость $v_{\text{гр}} = \frac{d\omega_{\text{оп}}}{dq}$ равна нулю.

Для длинных волн акустической ветви колебаний

$$\omega_{\text{ак}} \approx \frac{1}{4} \omega_0 \gamma a q = a \sqrt{\frac{\beta}{2(m_1 + m_2)}} q = v_{\text{зв}} q \quad (3.59)$$

частота колебания пропорциональна скорости звука, а фазовая и групповая скорости

$$v_{\text{ф}} = v_{\text{гр}} = v_{\text{зв}} \quad (3.60)$$

равны скорости звука.

Это значит, что при малых значениях q акустическая ветвь колебаний атомов решетки соответствует обычным звуковым волнам и скорость их распространения равна скорости звука в кристалле.

Для наименьшей длины волны $\lambda = 2a$, когда $q = \pi/a$, частота оптических колебаний будет минимальной:

$$\omega_{\text{оп}}(\pi/a) = (2\beta/m_2)^{1/2} = \omega_{\text{мин}}, \quad (3.61)$$

а частота акустических колебаний будет максимальной:

$$\omega_{\text{ак}}(\pi/a) = (2\beta/m_1)^{1/2} = \omega_{\text{ак макс}}. \quad (3.62)$$

Зависимость $\omega(q)$ для оптических и акустических ветвей колебаний одномерной решетки, имеющей в элементарной ячейке два атома, изображена на рис. 3.6. Во всем диапазоне изменений q $\omega_{\text{оп}} > \omega_{\text{ак}}$. Частота оптических колебаний незначительно уменьшается при увеличении q . Наибольшее отличие $\omega_{\text{оп}}$ от $\omega_{\text{ак}}$ имеет место при $q = 0$.

3.5. КОЛЕБАНИЯ АТОМОВ ТРЕХМЕРНОЙ РЕШЕТКИ

Установленная закономерность колебаний одномерной цепочки атомов позволяет высказать общие соображения о тепловом движении атомов трехмерного кристалла. Смещение атомов трехмерной решетки также можно описать посредством суперпозиции нормальных колебаний решетки. Каждое нормальное колебание представляет собой волну, направление распространения которой определяется волновым вектором q . Волна характеризуется длиной волны λ

и частотой ω_q . Волновой вектор $\mathbf{q}$ имеет столько разрешенных значений, сколько в кристалле имеется элементарных ячеек. При этом значения q меняются в пределах

$$-\pi \leq \mathbf{q} \cdot \mathbf{a}_i \leq \pi, \quad (3.63)$$

где $\mathbf{a}_i$ — постоянная решетки ($i = x, y, z$).

Объемный кристалл содержит $N = N_x N_y N_z$ элементарных ячеек, в каждой из которых может быть s атомов. Поскольку каждый атом обладает тремя степенями свободы, то кристалл характеризуется $3sN$ степенями свободы и в таком кристалле будут $3sN$ типов волн.

В трехмерном кристалле имеют место два типа упругих волн; продольные — волны сжатий и разрежений и поперечные — волны деформаций сдвига. Для первого типа направление распространения волны совпадает с направлением колебаний атомов кристалла. Поперечные волны, распространяющиеся в направлении, перпендикулярном плоскости колебания атомов, имеют два главных направления поляризации, так как поперечную волну можно разложить на две независимые поперечные волны, колебания в которых происходят по двум взаимно перпендикулярным направлениям.

Для объемного кристалла, содержащего s атомов в элементарной ячейке, должно быть $2s$ типа поперечных волн и s типов продольных волн, всего $3s$ типов волн, т. е. должно быть $3s$ различных ветвей колебаний. Другими словами, каждому значению $\mathbf{q}$ соответствует $3s$ колебаний, различающихся частотой и поляризацией. Три ветви колебаний являются акустическими. Для них в области длинных волн имеет место $v_{\text{ф}} = v_{\text{гр}} = v_{\text{зв}}$. При этом частота продольных колебаний больше частоты поперечных колебаний, так как скорость продольных упругих волн больше скорости поперечных волн. Остальные $(3s-3)$ ветви будут оптическими, у которых при $q \rightarrow 0$ $v_{\text{ф}} \rightarrow \infty$, а $v_{\text{гр}} \rightarrow 0$.

В общем случае дисперсионные кривые для волн с различной поляризацией отличаются друг от друга. Однако для изотропного кристалла две поперечные волны вырождены, им соответствуют одинаковые значения частоты ω_q . Вырождение различных ветвей колебания может иметь место и в анизотропном кристалле для направлений, характеризующихся высокой степенью симметрии.

На рис. 3.7 приведены измеренные колебательные спектры решеток германия, кремния и арсенида галлия. У них в элементарной ячейке находятся два атома, поэтому эти полупроводники имеют оптические и акустические ветви колебаний, каждая из которых содержит одну продольную и две поперечные ветви колебаний, являющиеся вырожденными.

Как следует из рис. 3.7, частота акустических колебаний стремится к нулю при $q \rightarrow 0$, что является типичным для акустической волны. Это имеет место для всех трех направлений колебаний — одного продольного LA и двух поперечных TA (вырожденные). При малых значениях волнового вектора $\mathbf{q}$ частота этих колебаний пропорциональна q для каждого направления распространения

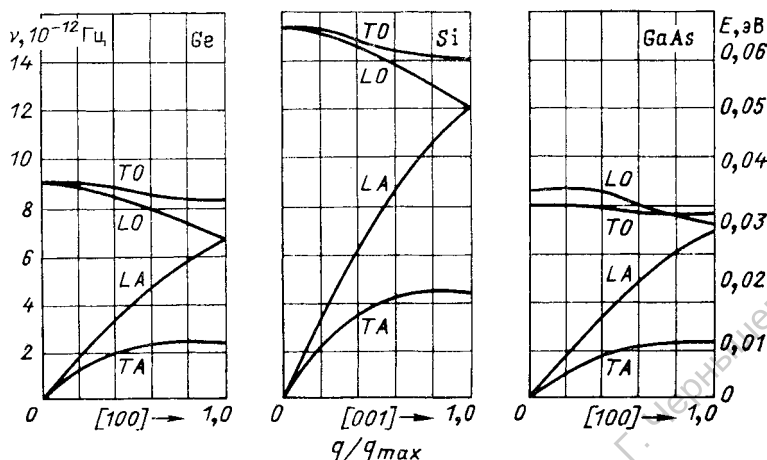


Рис. 3.7. Колебательные спектры решетки германия, кремния и арсенида галлия: ТА и LA — поперечные и продольные акустические колебания; TO и LO — поперечные и продольные оптические колебания

волны, а фазовая скорость постоянна и имеет два разных значения для LA и TA. Это также является характерным для акустических волн в кристалле.

Для всех трех направлений колебаний оптической ветви частота максимальна, когда $q \rightarrow 0$, и достигает 10^{13} Гц. При $q \rightarrow 0$ фазовая скорость стремится к бесконечности, а групповая — к нулю, что является характерным для оптической ветви колебаний. С увеличением q частота оптических колебаний решетки уменьшается незначительно, а поэтому можно считать, что $\omega_{\text{оп}} \approx \omega_0$, в то время как частота акустических колебаний резко возрастает. В области $q = \pi/a$, т. е. вблизи границы зоны Бриллюэна, различие между оптическими и акустическими колебаниями теряется.

Итак, колебание атомов кристалла может быть представлено как суперпозиция $3sN$ нормальных колебаний. Но каждое нормальное колебание с механической точки зрения представляет собой гармонический осциллятор. Следовательно, полная энергия колебаний кристалла равна сумме энергий колебаний $3sN$ не взаимодействующих между собой гармонических осцилляторов. С точки зрения квантовой механики qj -му осциллятору, колеблющемуся с частотой $\omega_j(\mathbf{q})$, нужно приписать энергию $E_{qj} = \hbar\omega_j(\mathbf{q}) \left[n_{qj} + \frac{1}{2} \right]$.

Таким образом, полная энергия колебаний кристалла

$$E = \sum_{qj} E_{qj} = \sum_{qj} \hbar\omega_j(\mathbf{q}) \left[n_{qj} + \frac{1}{2} \right] \quad (n_{qj} = 0, 1, 2, \dots). \quad (3.64)$$

Здесь $\mathbf{q}$ — волновой вектор, имеющий N различных значений; j — ветвь колебаний (всего имеется $3s$ ветвей колебаний или $3s$ типов колебаний).

Квантовый осциллятор с энергией $E_{qj} = \hbar \omega_j(\mathbf{q}) \left[n_{qj} + \frac{1}{2} \right]$ может изменять энергию на значение $\Delta E_{qj} = \hbar \omega_j(\mathbf{q}) \Delta n_{qj}$. При этом изменение квантового числа осциллятора согласно квантовой механике должно быть равно:

$$\Delta n_{qj} = \pm 1. \quad (3.65)$$

Квант энергии $\hbar \omega_j(\mathbf{q})$ носит название кванта энергии колебаний решетки или фонона. В том случае, если $\Delta n_{qj} = +1$, решетка переходит в более высокое энергетическое состояние. Поэтому переход $\Delta n_{qj} = +1$ можно назвать процессом поглощения фонона решеткой, а переход $\Delta n_{qj} = -1$ — процессом излучения фонона решеткой. Следовательно, фонон есть единичное квантовое возбуждение определенного нормального колебания.

По аналогии с фотоном — квантом электромагнитного поля — фонон можно рассматривать как носитель кванта энергии колебаний решетки, т. е. рассматривать как квазичастицу. Таким образом, колебания решетки в энергетическом отношении эквивалентны фононному газу. В этом случае $\Delta n_{qj} = +1$ будет означать рождение фонона, а $\Delta n_{qj} = -1$ — уничтожение фонона.

3.6. СТАТИСТИКА ФОНОНОВ

Для того чтобы определить среднее число фононов $\langle n_{qj} \rangle$ с энергией $\hbar \omega_j(\mathbf{q})$, существующих в данном нормальном колебании при температуре T , необходимо знание средней энергии осциллятора $\langle E_{qj} \rangle$.

Следовательно,

$$\langle n_{qj} \rangle = \frac{\langle E_{qj} \rangle}{\hbar \omega_j(\mathbf{q})}. \quad (3.66)$$

Пусть гармонический осциллятор в тепловом равновесии при температуре T находится в s -м квантовом состоянии с энергией E_s . Вероятность того, что осциллятор находится в этом состоянии, как доказывается в статистической физике, равна:

$$W_s = C e^{-E_s/kT}. \quad (3.67)$$

Коэффициент C определяется из условия нормировки вероятности

$$\sum_s W_s = 1, \quad (3.68)$$

из которого следует, что

$$C = \left(\sum_s e^{-E_s/kT} \right)^{-1}. \quad (3.69)$$

Средняя, или равновесная, энергия осциллятора равна сумме энергий E_s , помноженных на соответствующие вероятности W_s , т. е.

$$\langle E \rangle = \sum_{n=0}^{\infty} E_s W_s. \quad (3.70)$$

Учитывая равенства (3.68) и (3.69), выражение для равновесного значения энергии гармонического осциллятора запишем в виде

$$\langle E \rangle = \frac{\sum_{n=0}^{\infty} E_s e^{-E_s/kT}}{\sum_{n=0}^{\infty} e^{-E_s/kT}} = - \frac{d}{d \left(\frac{1}{kT} \right)} \ln \sum_{n=0}^{\infty} e^{-E_s/kT}. \quad (3.71)$$

Найдем значение суммы в (3.71), учитывая, что энергия гармонического осциллятора $E_s = \hbar\omega \left(n + \frac{1}{2} \right)$. Определить ее можно как сумму геометрической прогрессии:

$$\begin{aligned} \sum_{n=0}^{\infty} e^{-E_s/kT} &= \sum_{n=0}^{\infty} e^{-[\hbar\omega (n+1/2)]/kT} = \\ &= e^{-\hbar\omega/2kT} (1 + e^{-\hbar\omega/kT} + e^{-2\hbar\omega/kT} + \dots) = \frac{e^{-\hbar\omega/2kT}}{1 - e^{-\hbar\omega/kT}}. \end{aligned} \quad (3.72)$$

Если в (3.71) подставим (3.72) и проведем дифференцирование по аргументу $1/(kT)$, то получим:

$$\begin{aligned} \langle E \rangle &= - \frac{d}{d \left(\frac{1}{kT} \right)} \ln \frac{e^{-\hbar\omega/2kT}}{1 - e^{-\hbar\omega/kT}} = \frac{\hbar\omega}{2} + \frac{d}{d \left(\frac{1}{kT} \right)} \ln (1 - e^{-\hbar\omega/kT}) = \\ &= \frac{\hbar\omega}{2} + \frac{\hbar\omega e^{-\hbar\omega/kT}}{1 - e^{-\hbar\omega/kT}} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}. \end{aligned} \quad (3.73)$$

Член $\frac{1}{2}\hbar\omega$, не зависящий от температуры, называют нулевой энергией осциллятора. Поскольку во взаимодействии с электронами принимают участие только те нормальные колебания, которые возбуждены выше нулевого энергетического уровня, то нулевую энергию можно не учитывать. Поэтому будем считать, что средняя энергия гармонического осциллятора равна:

$$\langle E \rangle = \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}. \quad (3.74)$$

Тогда

$$\langle n \rangle = \frac{\langle E \rangle}{\hbar\omega} = \frac{1}{e^{\hbar\omega/kT} - 1}. \quad (3.75)$$

Из этого выражения следует, что фононы подчиняются статистике Бозе—Эйнштейна. $\langle n \rangle$ есть среднее число фононов в данном нормальном колебании, т. е. среднее число фононов (их концентрация) с энергией $\hbar\omega$ при тепловом равновесии в одном квантовом состоянии — ячейке фазового пространства объемом h^3 с энергией $\hbar\omega$.

При высоких температурах, когда $kT \gg \hbar\omega$, осциллятор находится в высоком энергетическом состоянии и число фононов в одной ячейке фазового пространства составляет:

$$\langle n \rangle \approx kT/\hbar\omega, \quad (3.76)$$

а средняя энергия осциллятора

$$\langle E \rangle = \langle n \rangle \hbar \omega = kT, \quad (3.77)$$

т. е. соответствует классической статистике.

При низких температурах, когда $kT < \hbar \omega$,

$$\langle n \rangle = e^{-\hbar \omega / kT}, \quad (3.78)$$

а средняя энергия осциллятора

$$\langle E \rangle = \hbar \omega e^{-\hbar \omega / kT}. \quad (3.79)$$

Итак, число фононов, находящихся в одном квантовом состоянии, зависит от температуры. Кроме того, при взаимодействии с электронами или друг с другом фононы возникают и исчезают. Поэтому фононы называют квазичастицами, поскольку они ведут себя не как газ обыкновенных частиц (атомы, электроны), для которых объем, температура и число частиц являются независимыми.

В твердом теле возможны акустические и оптические колебания, но каждое колебание решетки в энергетическом отношении эквивалентно фонону. Поэтому можно говорить об акустических и оптических фононах. Их среднее число, приходящееся на одно квантовое состояние, согласно (3.75) зависит от энергии фононов $\hbar \omega_j(\mathbf{q})$ и температуры T .

Во всем диапазоне изменения q от 0 до π/a $\omega_{j \text{ оп}} > \omega_{j \text{ ак}}$, поэтому энергия оптических фононов больше энергии акустических фононов (см. рис. 3.7). При малых значениях q энергия акустических фононов, равная

$$E_{\text{ак}, \Phi} = \hbar \omega_{j \text{ ак}} = \hbar v_{j \text{ зв}} q,$$

значительно меньше энергии оптических фононов, для которых ее значение максимально: $E_{\text{оп}, \Phi} = \hbar \omega_{0j}$.

При низких температурах, когда $kT \ll \hbar \omega_{0j}$, оптические фононы отсутствуют, возбуждаются только акустические фононы, а число их в одном квантовом состоянии определяется равенством (3.78). При высоких температурах, когда $kT \gg \hbar \omega_{0j}$, кроме акустических фононов возбуждается большое количество и оптических фононов.

3.7. ТЕПЛОЕМКОСТЬ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ

По определению теплоемкость твердого тела

$$C = \frac{\partial E}{\partial T}. \quad (3.80)$$

Поскольку гармонические осцилляторы не взаимодействуют, то полная энергия кристалла в термодинамическом равновесии E равна сумме средних значений энергии всех осцилляторов. При этом суммирование нужно провести с учетом волнового вектора $\mathbf{q}$, имеющего N различных значений, и типов колебаний, которых может быть 3s. Кроме того, нужно учесть нулевую энергию

$E_0 = \sum_{\mathbf{q}j} \frac{1}{2} \hbar \omega_j(\mathbf{q})$. Поэтому на основании (3.74) энергия колебаний решетки может быть записана в виде

$$E = E_0 + \sum_{j=1}^{3s} \sum_{\mathbf{q}} \frac{\hbar \omega_j(\mathbf{q})}{e^{\hbar \omega_j(\mathbf{q})/kT} - 1} = E_0 + \langle E_A \rangle + \langle E_O \rangle. \quad (3.81)$$

Здесь равновесное значение энергии акустических колебаний решетки

$$\langle E_A \rangle = \sum_{j=1}^3 \sum_{\mathbf{q}} \frac{\hbar \omega_j(\mathbf{q})}{e^{\hbar \omega_j(\mathbf{q})/kT} - 1}, \quad (3.82)$$

а равновесное значение энергии оптических колебаний решетки

$$\langle E_O \rangle = \sum_{j=4}^{3s} \sum_{\mathbf{q}} \frac{\hbar \omega_j(\mathbf{q})}{e^{\hbar \omega_j(\mathbf{q})/kT} - 1}. \quad (3.83)$$

Определим $\langle E_A \rangle$. Так как волновой вектор $\mathbf{q}$ меняется квазинепрерывно от 0 до π/a , имея N значений, равных числу элементарных ячеек в кристалле, то и частота колебаний $\omega_j(\mathbf{q})$ также меняется квазинепрерывно. В акустической ветви колебаний она меняется от 0 до $\omega_{A \max}$. Поскольку N велико, то в (3.82) сумму по $\omega_j(\mathbf{q})$, выражающую энергию акустических колебаний $\langle E_A \rangle$, можно заменить интегралом:

$$\langle E_A \rangle = \int_0^{\omega_{A \max}} dZ \hbar \omega \langle n \rangle = \int_0^{\omega_{A \max}} dZ \frac{\hbar \omega}{e^{\hbar \omega/kT} - 1}, \quad (3.84)$$

где dZ — число нормальных колебаний в интервале от q до $q + dq$, а $\omega_{A \max}$ соответствует $q_{\max}$.

Для определения числа нормальных колебаний dZ во всех трех акустических ветвях поступим следующим образом. В $\mathbf{q}$ -пространстве выделим слой, заключенный между сферическими поверхностями радиусом q и $q + dq$. Объем этого слоя (заштрихованная область на рис. 3.8) $dV_q = 4\pi q^2 dq$. Объем элементарной ячейки зоны Бриллюэна кубического кристалла ΔV_B в $\mathbf{q}$ -пространстве согласно (3.21) равен:

$$\Delta V_B = dq_x dq_y dq_z = \frac{2\pi}{aN_x} \frac{2\pi}{aN_y} \frac{2\pi}{aN_z} = \frac{8\pi^3}{V}, \quad (3.85)$$

где $V = a^3 N_x N_y N_z = a^3 N$ — объем кристалла; $N = N_x N_y N_z$ — число элементарных ячеек в кристалле.

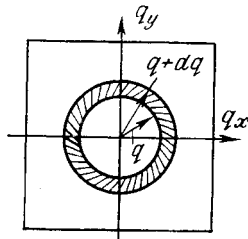


Рис. 3.8. Объем слоя, заключенного между двумя поверхностями

Каждой элементарной ячейке зоны Бриллюэна соответствует одно возможное колебательное состояние кристалла (т. е. одно нормальное колебание). Следовательно, в одной акустической ветви на интервал dq приходится нормальных колебаний в количестве

$$dZ = \frac{dV_{\mathbf{q}}}{\Delta V_B} = V \frac{q^2 dq}{2\pi^3}. \quad (3.86)$$

В соответствии с теорией теплоемкости по Дебаю будем считать, что для трех акустических ветвей для всей области изменения q (от 0 до π/a) имеет место закон дисперсии

$$\omega_j(\mathbf{q}) = v_{j\text{зв}} q \quad (j = 1, 2, 3). \quad (3.87)$$

Надо отметить, что в этом случае кристаллическое тело рассматривается как изотропный континуум, в котором отсутствует дисперсия. Теперь с учетом (3.87) число колебаний в одной акустической ветви в интервале dq , а следовательно, в интервале частот от ω до $\omega + d\omega$ равно:

$$dZ = V \frac{q^2 dq}{2\pi^3} = V \frac{\omega^2 d\omega}{2\pi^2 v_{j\text{зв}}^3}. \quad (3.88)$$

Найдем теперь плотность состояний колебаний решетки $g(\omega)$, т. е. число нормальных колебаний, соответствующих единичному интервалу частоты кристалла единичного объема; $g(\omega)$ также называется функцией распределения частот. Согласно (3.88) она имеет вид:

$$g(\omega) = \frac{dZ}{V d\omega} = \frac{\omega^2}{2\pi^2 v_{j\text{зв}}^3}. \quad (3.89)$$

Для трех возможных типов акустических колебаний — одного продольного со скоростью $v_{l\text{зв}}$ и двух поперечных со скоростью $v_{t\text{зв}}$ — плотность состояний колебаний в интервале ω , $\omega + d\omega$ равна:

$$G(\omega) = \frac{\omega^2}{2\pi^2 v_{l\text{зв}}^3} + \frac{2\omega^2}{2\pi^2 v_{t\text{зв}}^3} = \frac{3\omega^2}{2\pi^2 v_0^3}, \quad (3.90)$$

где v_0 , определяемая условием

$$\frac{1}{v_0^3} = \frac{1}{3} \left[\frac{1}{v_{l\text{зв}}^3} + \frac{2}{v_{t\text{зв}}^3} \right], \quad (3.91)$$

есть скорость звука, усредненная по направлениям и типам колебаний.

В этом случае для кристалла объемом V энергия акустических колебаний (3.84) запишется в виде

$$\langle E_A \rangle = \int_0^{\omega_{A\text{max}}} VG(\omega) \frac{\hbar\omega d\omega}{e^{\hbar\omega/kT} - 1} = \frac{3V\hbar}{2\pi^2 v_0^3} \int_0^{\omega_{A\text{max}}} \frac{\omega^3 d\omega}{e^{\hbar\omega/kT} - 1}. \quad (3.92)$$

Введем новые переменные

$$x = \hbar\omega/kT \text{ и } \hbar\omega_{A \max}/k = \theta,$$

тогда

$$x_{\max} = \hbar\omega_{A \max}/kT = \theta/T. \quad (3.93)$$

В этом случае

$$\langle E_A \rangle = \frac{3V\hbar}{2\pi^2 v_{0 \text{ зв}}^3} \left(\frac{kT}{\hbar} \right)^4 \int_0^{x_{\max}} \frac{x^3 dx}{e^x - 1} = \frac{3V\hbar}{2\pi^2 v_{0 \text{ зв}}^3} \left(\frac{kT}{\hbar} \right)^4 \int_0^{\theta/T} \frac{x^3 dx}{e^x - 1}. \quad (3.94)$$

Максимальная частота $\omega_{A \max}$ может быть определена из условия, что полное число акустических колебаний равно $3N$. Тогда

$$\int_0^{\omega_{A \max}} VG(\omega) d\omega = \frac{3V}{2\pi^2 v_{0 \text{ зв}}^3} \int_0^{\omega_{A \max}} \omega^2 d\omega = \frac{V\omega_{A \max}^3}{2\pi^2 v_{0 \text{ зв}}^3} = 3N, \quad (3.95)$$

откуда

$$\omega_{A \max} = v_{0 \text{ зв}} (6\pi^2 N/V)^{1/3}. \quad (3.96)$$

Параметр θ , равный

$$\theta = \frac{\hbar\omega_{A \max}}{k} = \frac{\hbar v_{0 \text{ зв}}}{k} \left(\frac{6\pi^2 N}{V} \right)^{1/3}, \quad (3.97)$$

называют *характеристической температурой твердого тела* или температурой Дебая. Ее физический смысл состоит в том, что $k\theta$ — это энергия акустических фононов в теле, для которых $q_{\max} = \pi/a$. Длина волны таких фононов $\lambda = 2a = 6 \cdot 10^{-8}$ см, где $a = 3 \cdot 10^{-8}$ см — постоянная решетки. Так как $\omega_{A \max} \approx v_{\text{зв}} a \approx 10^{13} \text{ с}^{-1}$, $\hbar = 1,05 \cdot 10^{-34} \text{ Дж} \cdot \text{с}$, $k = 1,38 \times 10^{-23} \text{ Дж/К}$, то $\theta \approx 100 \text{ К}$. Для большинства твердых тел θ лежит в пределах 100—300 К (табл. 3.1).

Таблица 3.1. Характеристическая температура Дебая θ некоторых полупроводников

Вещество	θ , К	Вещество	θ , К
Au	170	Ge	290
Cd	160	NaCl	280
Cu	315	KCl	230
Sn (серое)	260	KBr	177

Учитывая (3.97), коэффициент перед интегралом в выражении (3.94) можно записать в виде

$$\frac{3V\hbar}{2\pi^2 v_{0 \text{ зв}}^3} \left(\frac{kT}{\hbar} \right)^4 = 9kTN \left(\frac{T}{\theta} \right)^3, \quad (3.98)$$

поэтому

$$\langle E_A \rangle = \frac{9NkT}{(\theta/T)^3} \int_0^{\theta/T} \frac{x^3 dx}{e^x - 1} = 3NkTD \left(\frac{\theta}{T} \right), \quad (3.99)$$

где функция Дебая

$$D \left(\frac{\theta}{T} \right) = \frac{3}{(\theta/T)^3} \int_0^{\theta/T} \frac{x^3 dx}{e^x - 1}. \quad (3.100)$$

Определим теперь $\langle E_O \rangle$. Частоты оптических колебаний мало меняются во всем диапазоне изменений волнового вектора $\mathbf{q}$. Можно считать их равными предельным значениям $\omega_j(\mathbf{q}) = \omega_{0j}$. Введем температуру Дебая, соответствующую предельной частоте колебаний данной оптической ветви:

$$\theta_{0j} = \frac{\hbar \omega_{0j}}{k}. \quad (3.101)$$

Для большинства веществ θ_0 лежит в пределах 100—800 К.

В табл. 3.2 приведены значения энергии продольных оптических фононов $\hbar \omega_0$ и соответствующие ей температуры Дебая.

Таблица 3.2. Ширина запрещенной зоны, максимальная частота продольных оптических фононов ω_0 , их энергия $\hbar \omega_0$ и температура Дебая θ_0 для некоторых полупроводников

Полупроводники	Ширина запрещенной зоны (300 К), эВ	$\omega_0 \cdot 10^{-13}$, Гц	$\hbar \omega_0$, эВ	θ_0 , К
Si	1,08	9,45	0,063	730
Ge	0,66	4,8	0,037	430
GaAs	1,35	5,35	0,0355	408
GaSb	0,67	4,4	0,029	336
InAs	0,35	4,4	0,029	334
InSb	0,17	3,5	0,0225	264

Используя (3.101), энергию оптических колебаний $\langle E_O \rangle$ можно представить в виде

$$\langle E_O \rangle = \sum_{j=4}^{3s} \sum_{\mathbf{q}} \frac{\hbar \omega_j(\mathbf{q})}{e^{\hbar \omega_j(\mathbf{q})/kT} - 1} = NkT \sum_{j=4}^{3s} \frac{\theta_{0j}/T}{e^{\theta_{0j}/T} - 1}, \quad (3.102)$$

так как суммирование по $\mathbf{q}$ для каждой оптической ветви колебаний эквивалентно умножению на N .

Следовательно, энергия колебаний решетки

$$E = E_0 + \langle E_A \rangle + \langle E_O \rangle = E_0 + NkT \left\{ 3D \left(\frac{\theta}{T} \right) + \sum_{j=4}^{3s} \frac{\theta_{0j}/T}{e^{\theta_{0j}/T} - 1} \right\}. \quad (3.103)$$

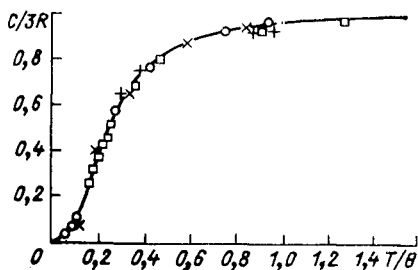


Рис. 3.9. Теплоемкость некоторых металлов:

× — Al, $\Theta = 396$ К; ○ — Cu, $\Theta = 309$ К;
□ — Ag, $\Theta = 215$ К; + — Pb, $\Theta = 95$ К

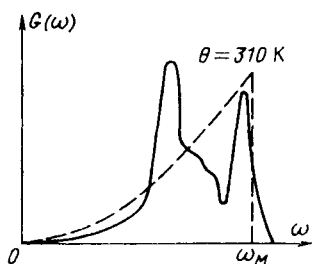


Рис. 3.10. Плотность колебаний для решетки вольфрама (сплошная кривая) и дебаевская кривая (пунктирная кривая)

В общем случае выражение для теплоемкости $C = \frac{\partial E}{\partial T}$ получается достаточно сложным. Рассмотрим область высоких и низких температур. Для области высоких температур, когда $T \gg \theta_{0j}$ и тем более $T \gg \theta$ (так как $\theta_{0j} > \theta$),

$$e^x \approx 1 + x \text{ и } 3D\left(\frac{\theta}{T}\right) = 9 \left(\frac{T}{\theta}\right)^3 \int_0^{\theta/T} x^2 dx = 3,$$

а сумма

$$\sum_{j=1}^{3s} \frac{\theta_{0j}/T}{e^{\theta_{0j}/T} - 1} = \sum_{j=1}^{3s} 1 = 3s - 3, \quad (3.104)$$

поэтому

$$E = E_0 + NkT [3 + (3s - 3)] = E_0 + 3sNkT = E_0 + 3RT, \quad (3.105)$$

где

$$R = skN = 8,31 \text{ Дж/(моль} \cdot \text{К)}. \quad (3.106)$$

Теплоемкость решетки

$$C = \frac{\partial E}{\partial T} = 3skN = 3R. \quad (3.107)$$

Следовательно, при высоких температурах теплоемкость твердого тела не зависит от температуры (рис. 3.9). Для монокристаллического твердого тела $C = 24,93 \text{ Дж/(моль} \cdot \text{К)}$.

При очень низких температурах, когда $T < \theta_{0j}$, оптические колебания не возбуждаются, а поэтому не дают вклада в энергию колебаний решетки. Для $T \ll \theta$ верхний предел интеграла в (3.100) можно положить равным ∞ . Так как

$$\int_0^{\infty} \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15}, \quad (3.108)$$

то

$$E = E_0 + \frac{3\pi^4 NkT^4}{5\theta^3}, \quad (3.109)$$

а теплоемкость решетки

$$C = \frac{12\pi^4 Nk}{5} \left(\frac{T}{\theta} \right)^3. \quad (3.110)$$

Таким образом, при низких температурах теплоемкость твердого тела пропорциональна T^3 . Закон $C \sim T^3$, полученный из теории Дебая, оправдывается только в области нескольких градусов вблизи абсолютного нуля (рис. 3.9). Отступление истинной теплоемкости от дебаевского значения (3.110) определяется тем, что истинная плотность колебаний $G(\omega)$ для акустических ветвей в кристалле сильно отличается от зависимости (3.90). Это положение иллюстрирует рис. 3.10, где представлены рассчитанная функция $G(\omega)$ для решетки вольфрама (сплошная кривая) и дебаевская функция $G(\omega)$ (пунктирная кривая) для температуры Дебая вольфрама, равной 310 К.

3.8. ТЕРМИЧЕСКОЕ РАСШИРЕНИЕ И ТЕПЛОВОЕ СОПРОТИВЛЕНИЕ ТВЕРДОГО ТЕЛА

При определении потенциальной энергии в уравнении (3.37) был учтен только квадратичный член смещения, так называемый гармонический член. Допустим, что это смещение равно x , тогда потенциальная энергия $U = \beta x^2$.

Согласно Больцману вероятность отклонения атома от положения равновесия на величину x равна:

$$f(x) = Ae^{-U/kT}. \quad (3.111)$$

В этом случае расстояние между атомами, совершающими гармоническое колебание, не изменяется, так как их среднее смещение, которое может быть определено по формуле

$$\bar{x} = \frac{\int_{-\infty}^{\infty} xf(x) dx}{\int_{-\infty}^{\infty} f(x) dx} = \frac{\int_{-\infty}^{\infty} xe^{-\beta x^2/kT} dx}{\int_{-\infty}^{\infty} e^{-\beta x^2/kT} dx} = 0, \quad (3.112)$$

равно нулю (интеграл числителя дроби равен нулю ввиду нечетности подынтегральной функции).

Реальные тела испытывают тепловое расширение. Его можно объяснить, если учесть кубические члены разложения потенциальной энергии по смещениям атомов. Расчет ангармонических членов сил взаимодействия атомов твердого тела представляет собой очень сложную задачу. Однако рассмотрение простой модели двух взаимодействующих атомов позволяет понять причины теплового расширения. Кроме того, эта модель дает возможность определить пра-

вильный порядок значения температурного коэффициента линейного расширения.

Предположим, что два атома при смещении на расстояние x от положения их равновесия в решетке взаимодействуют друг с другом с силой

$$F = -\frac{dU}{dx} = -2\beta x + 3\gamma x^2 \quad (3.113)$$

и потенциальной энергией

$$U = \beta x^2 - \gamma x^3. \quad (3.114)$$

Если в уравнение (3.111) подставить значение U из (3.114), то в случае малых смещений экспоненту, соответствующую ангармоническому члену, можно разложить в ряд, т. е. $e^{\gamma x^3/kT} = 1 + \gamma x^3/kT$. В результате получим приближенно:

$$f(x) \approx A e^{-\beta x^2/kT} \left(1 + \frac{\gamma x^3}{kT}\right). \quad (3.115)$$

Среднее отклонение атома от положения равновесия

$$\bar{x} = \frac{\int_{-\infty}^{\infty} e^{-\beta x^2/kT} (1 + \gamma x^4/kT) dx}{\int_{-\infty}^{\infty} e^{-\beta x^2/kT} (1 + \gamma x^3/kT) dx} = \frac{\frac{\gamma}{kT} \frac{3V\pi}{4} (kT/\beta)^{5/2}}{(\pi kT/\beta)^{1/2}} = \frac{3\gamma kT}{4\beta^2}, \quad (3.116)$$

где интегралы от первого слагаемого в числителе дроби, содержащего множитель x , и от второго слагаемого в знаменателе дроби, содержащего x^3 , равны нулю ввиду нечетности подынтегральной функции.

Как известно, линейное тепловое расширение определяется формулой

$$l_t = l_0 (1 + \alpha t), \quad (3.117)$$

поэтому в нашем случае температурный коэффициент линейного расширения α , определяемый как удлинение на единицу длины и на 1°C , равен:

$$\alpha = \bar{x}/aT = 3k\gamma/4a\beta^2, \quad (3.118)$$

где a — параметр решетки.

Таким образом, температурный коэффициент линейного расширения твердого тела есть постоянная величина; он пропорционален коэффициенту ангармоничности и при $\gamma = 0$ равен нулю. Значения коэффициентов линейного расширения для некоторых веществ приведены в табл. 3.3.

Ангармонизм колебаний решетки позволяет также объяснить тепловое сопротивление твердого тела. При гармонических колебаниях решетки нормальные координаты являются независимыми, т. е. имеет место принцип линейных суперпозиций, согласно которому

Т а б л и ц а 3.3. Температурные коэффициенты линейного расширения при комнатной температуре

Вещество	$\alpha \cdot 10^6, \text{K}^{-1}$	Вещество	$\alpha \cdot 10^6, \text{K}^{-1}$
Au	14	Pt	9
Li	56	CsCl	50
Ni	13	AgBr	400

упругие волны распространяются в кристалле независимо, не рассеиваясь друг на друге. Поскольку каждой нормальной координате, представляющей собой гармонический осциллятор, соответствует определенный набор фононов, то в гармоническом приближении фононы не взаимодействуют, т. е. не сталкиваются друг с другом. В этом случае тепловой поток — энергия, переносимая фононами, — распространяется со скоростью звука ($v_{\text{гр}} = v_{\text{зв}}$), поэтому тепловое сопротивление равно нулю. Учет ангармонических членов в потенциальной энергии, т. е. учет членов третьей степени в смещении атомов (3.114), делает нормальные координаты решетки, а следовательно, и осцилляторы ангармоническими и взаимно зависимыми. Ангармонизм осцилляторов и обуславливает рассеяние фононов на фононах. При этом в процессе столкновения участвуют три фонона, в результате чего либо два фонона превращаются в один, либо один фонон распадается на два. Рассеяние фононов на фононах, сопровождающееся рождением и исчезновением фононов, и приводит к появлению теплового сопротивления в твердом теле. С ростом температуры ангармонизм колебаний усиливается, так как возрастает амплитуда колебаний, поэтому увеличивается и тепловое сопротивление.

ГЛАВА ЧЕТВЕРТАЯ

СТАТИСТИКА ЭЛЕКТРОНОВ И ДЫРОК В ПОЛУПРОВОДНИКАХ

4.1. ПЛОТНОСТЬ КВАНТОВЫХ СОСТОЯНИЙ

Важнейшая задача статистической физики состоит в определении числа частиц, энергия которых лежит в определенном интервале. Для ее решения необходимо знать число квантовых состояний и вероятность нахождения частиц в этих состояниях. Следовательно, для определения концентрации носителей заряда в полупроводнике необходимо знать фактическое число состояний, занятых электронами и дырками.

Пусть в кристалле единичного объема в интервале энергий от E до $E + dE$ имеется dZ квантовых состояний (с учетом спина). Обозначим через $N(E)$ плотность состояний, т. е. число состояний

в единичном интервале энергии для единичного объема кристалла. Тогда согласно определению

$$N(E) = \frac{dZ}{dE}. \quad (4.1)$$

Если вероятность заполнения электроном состояния с энергией E равна $f(E, T)$, то число электронов dn , находящихся в состояниях dZ , составит величину

$$dn = f(E, T) dZ = f(E, T) N(E) dE. \quad (4.2)$$

Соответственно количество электронов, для которых возможный интервал энергии лежит в пределах от E_1 до E_2 , будет равно:

$$n = \int_{E_1}^{E_2} f(E, T) N(E) dE. \quad (4.3)$$

Найдем выражение для плотности квантовых состояний в случае, когда поверхности равной энергии зоны проводимости и валентной зоны являются сферами. Определим плотность состояний у нижнего края зоны проводимости. Согласно (2.139) энергия электронов у дна зоны может быть записана в виде

$$E = E_c + \frac{p^2}{2m_n^*}, \quad (4.4)$$

где $E_c = E(\mathbf{p}_0)$ — энергия электрона на дне зоны проводимости; m_n^* — эффективная масса электрона.

Выделим шаровой слой, заключенный между двумя изоэнергетическими поверхностями, соответствующими энергии $E(\mathbf{p}) = \text{const}$ и $E(\mathbf{p}) + dE = \text{const}$ (рис. 4.1). Объем этого слоя

$$dV_p = 4\pi p^2 dp. \quad (4.5)$$

Объем элементарной ячейки зоны Бриллюэна кристалла единичного объема в $\mathbf{p}$ -пространстве на основании (2.79) равен h^3 . В каждой ячейке могут находиться два электрона с противоположно направленными спинами. С учетом этого количество квантовых состояний в объеме dV_p равно:

$$dZ = 2 \frac{dV_p}{h^3} = \frac{8\pi}{h^3} p^2 dp. \quad (4.6)$$

Исходя из (4.4), имеем

$$p^2 = 2m_n^*(E - E_c), \quad (4.7)$$

откуда

$$p = (2m_n^*)^{1/2} (E - E_c)^{1/2}$$

и

$$dp = \frac{1}{2} (2m_n^*)^{1/2} (E - E_c)^{-1/2} dE. \quad (4.8)$$

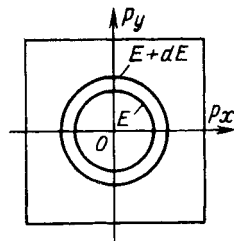


Рис. 4.1. Объем слоя в зоне Бриллюэна, заключенного между двумя изоэнергетическими поверхностями E и $E+dE$

Подставив (4.6), (4.7) и (4.8) в соотношение (4.1), получим выражение для плотности квантовых состояний у дна зоны проводимости, обладающей сферической симметрией:

$$N(E) = 4\pi (2m_p^*/\hbar^2)^{3/2} (E - E_c)^{1/2}. \quad (4.9)$$

Определим плотность состояний вблизи верхнего края валентной зоны. С этой целью воспользуемся (2.139) и (2.144) для энергии дырок:

$$E = E_v - \frac{p^2}{2m_p^*}. \quad (4.10)$$

Здесь $E_v = E(p_0)$ — энергия дырки у потолка валентной зоны; m_p^* — эффективная масса дырки. Из расчетов, аналогичных проведенным выше, будем иметь:

$$N(E) = 4\pi (2m_p^*/\hbar^2)^{3/2} (E_v - E)^{1/2}. \quad (4.11)$$

Таким образом, если энергия носителей заряда является квадратичной функцией квазиимпульса, то плотность состояний $N(E)$ имеет зависимость от энергии вида $(E - E_c)^{1/2}$ или $(E_v - E)^{1/2}$. Следовательно, выражения (4.9) и (4.11) справедливы только для состояний вблизи экстремума энергии, т. е. у дна и потолка зоны.

Подсчитаем теперь плотность квантовых состояний для случая сложной зоны проводимости, когда зависимость $E(p)$ у дна зоны описывается соотношением вида (2.132). Поскольку для дна зоны $p_{0x} = p_{0y} = p_{0z} = 0$, а $E(p_0) = E_c$, то

$$E = E_c + \frac{p_x^2}{2m_1} + \frac{p_y^2}{2m_2} + \frac{p_z^2}{2m_3}, \quad (4.12)$$

где $1/m_1, 1/m_2, 1/m_3$ — диагональные компоненты тензора обратной эффективной массы.

Изоэнергетические поверхности в этом случае представляют собой эллипсоиды, уравнение которых в канонической форме имеет вид:

$$\frac{p_x^2}{a^2} + \frac{p_y^2}{b^2} + \frac{p_z^2}{c^2} = 1, \quad (4.13)$$

где полуоси эллипсоида

$$\left. \begin{aligned} a &= [2m_1(E - E_c)]^{1/2}; \\ b &= [2m_2(E - E_c)]^{1/2}; \\ c &= [2m_3(E - E_c)]^{1/2}. \end{aligned} \right\} \quad (4.14)$$

Объем эллипсоида с полуосями a, b, c равен:

$$V_p = \frac{4\pi}{3} abc = \frac{8\pi}{3} (2m_1 m_2 m_3)^{1/2} (E - E_c)^{3/2}. \quad (4.15)$$

Объем слоя, заключенного между эллипсоидами равной энергии $E = \text{const}$ и $E + dE = \text{const}$, будет:

$$dV_p = 4\pi (2m_1 m_2 m_3)^{1/2} (E - E_c)^{1/2} dE. \quad (4.16)$$

В объеме dV_p с учетом спина заключено следующее количество квантовых состояний:

$$dZ = 2 \frac{dV_p}{h^3} = 4\pi \left(\frac{2}{h^2} \right)^{3/2} (m_1 m_2 m_3)^{1/2} (E - E_c)^{1/2} dE, \quad (4.17)$$

поэтому выражение для плотности состояний у дна сложной зоны проводимости примет вид:

$$N(E) = 4\pi (2/h^2)^{3/2} (m_1 m_2 m_3)^{1/2} (E - E_c)^{1/2}. \quad (4.18)$$

Его можно свести к виду, аналогичному выражению (4.9), если положить:

$$m_1 m_2 m_3 = m_{дн}^{*3}, \quad (4.19)$$

где $m_{дн}^*$ называют эффективной массой плотности состояний для электронов.

При этом, как и для простой зоны, получим

$$N(E) = 4\pi (2m_{дн}^*/h^2)^{3/2} (E - E_c)^{1/2}. \quad (4.20)$$

Таким образом, эллипсоидальную поверхность равной энергии зоны проводимости можно заменить сферой, если ввести эффективную массу плотности состояний для электронов, которая вычисляется по (4.19).

Определим эффективные массы плотности состояний для электронов и дырок в кремнии и германии.

К р е м н и й. Как следует из рис. 2.21, абсолютный минимум энергии в зоне проводимости кремния расположен в направлении $[100]$ на расстоянии от точки $k = 0$, равном примерно 75 % величины $2\pi/a$. Всего в первой зоне Бриллюэна располагается шесть минимумов, и для энергии, несколько превышающей величину E_c , имеется шесть эллипсоидов равной энергии, вытянутых вдоль главных кристаллографических осей (см. рис. 2.22). В соответствии с этим плотность состояний, которая выражается для одного эллипсоида равенством (4.18), увеличится в 6 раз. И если учесть, что для кремния $m_1 = m_2$, то

$$N(E) = 6 \cdot 4\pi (2/h^2) (m_1^2 m_3)^{1/2} (E - E_c)^{1/2}, \quad (4.21)$$

а эффективная масса плотности состояний для электронов с учетом значений $m_1 = 0,19m_0$ и $m_3 = 0,92m_0$ будет:

$$m_{дн}^* = 6^{2/3} (m_1^2 m_3)^{1/3} = 1,08m_0. \quad (4.22)$$

Следовательно, у кремния все шесть эллипсоидов изоэнергетической поверхности зоны проводимости можно заменить одной сферической поверхностью с эффективной массой плотности состояний для электронов, равной $1,08m_0$.

В точке $k = 0$ валентная зона кремния дважды вырождена и поверхности равной энергии имеют довольно сложный вид, описываемый формулой (2.153). Однако их можно аппроксимировать двумя сферическими поверхностями вида (2.154), которым соответствуют эффективные массы тяжелых $m_{рл}^*$ и легких дырок $m_{рл}^*$.

В этом случае плотность состояний будет определяться суммой плотности состояний в зонах тяжелых и легких дырок:

$$N(E) = 4\pi (2/h^2)^{3/2} (m_{pt}^{*3/2} + m_{pl}^{*3/2}) (E_v - E)^{1/2}. \quad (4.23)$$

Изоэнергетические поверхности обеих зон можно заменить одной приведенной сферой с плотностью состояний

$$N(E) = 4\pi (2m_{dp}^*/h^2)^{3/2} (E_v - E)^{1/2}, \quad (4.24)$$

для которой эффективная масса плотности состояний для дырок равна:

$$m_{dp}^* = (m_{pt}^{*3/2} + m_{pl}^{*3/2})^{2/3} = 0,56m_0, \quad (4.25)$$

так как $m_{pt}^* = 0,56m_0$, а $m_{pl} = 0,16m_0$.

Германий. В первой зоне Бриллюэна германия имеется восемь минимумов энергии, расположенных в точках $[1/2, 1/2, 1/2]$. Соответствующие поверхности равной энергии изображены на рис. 2.22, где показано восемь эллипсоидов энергии, рассеченных пополам гранями зоны Бриллюэна. Две противоположные половины эллипсоидов рис. 2.22 могут рассматриваться как одна целая эллипсоидальная поверхность. Таким образом, в основной зоне Бриллюэна германия расположено четыре полных эллипсоида, для которых $m_1 = m_2$. Следовательно, плотность состояний у потолка зоны проводимости германия будет определяться выражением (4.20) с эффективной массой плотности состояний для электронов, равной:

$$m_{dn}^* = (4^2 m_1^2 m_3)^{1/3} = 0,56m_0, \quad (4.26)$$

так как $m_1 = 0,082m_0$, а $m_3 = 1,59m_0$.

Структура валентной зоны германия подобна структуре зоны кремния (см. рис. 2.21). Эффективная масса плотности состояний для дырок, характеризующая плотность квантовых состояний у потолка валентной зоны германия, будет:

$$m_{dp}^* = (m_{pt}^{*3/2} + m_{pl}^{*3/2})^{2/3} = 0,35m_0, \quad (4.27)$$

так как $m_{pt}^* = 0,33m_0$; $m_{pl} = 0,04m_0$.

4.2. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ ФЕРМИ—ДИРАКА

Для определения числа частиц, имеющих энергию в заданном интервале, помимо плотности квантовых состояний $N(E)$ необходимо знать вероятность того, что данное состояние с энергией E занято частицей, т. е. нужно знать функцию распределения $f(E)$. В условиях теплового равновесия для частиц с полужелым спином, подчиняющихся принципу Паули, справедливо распределение Ферми—Дирака

$$f_0(E) = \frac{1}{e^{(E-F)/kT} + 1}, \quad (4.28)$$

где k — постоянная Больцмана; T — абсолютная температура; F — энергия Ферми или электрохимический потенциал, т. е. работа,

которую необходимо затратить для изменения числа частиц в системе на единицу при условии постоянства объема и температуры.

Рассмотрим вид функции распределения Ферми—Дирака при различных температурах. Из (4.28) следует, что в случае $T = 0$ в интервале энергии $0 \leq E < F$ имеем $f_0 = 1$ и $f_0 = 0$ для $E > F$. Это означает, что все квантовые состояния с энергией, меньшей энергии Ферми, заняты электронами, а уровни, лежащие выше уровня Ферми, полностью свободны, не заняты электронами. Следовательно, энергия Ферми есть максимально возможная энергия электронов в металле при температуре абсолютного нуля.

Рассмотрим случай, когда $T > 0$. Из (4.28) для значения энергии, равной значению энергии Ферми ($E = F$), имеем $f_0 = 1/2$. Таким образом, уровень Ферми есть энергетический уровень, вероятность заполнения которого при температуре, отличной от абсолютного нуля, равна 0,5. При $T > 0$ часть электронов в результате теплового движения перейдет в состояния с энергией, большей энергии Ферми ($E > F$), и соответственно часть состояний, находящихся ниже уровня Ферми, окажется свободной. В этом случае число частиц, перешедших на более высокие энергетические уровни, будет равно количеству образовавшихся свободных состояний в области $E < F$.

Произведем оценку области изменения функции распределения $f_0(E)$ для случая $T > 0$. Для этого подсчитаем $f_0(E)$ для разных значений энергии. Для энергий, отличающихся от F на $\pm kT$, значение $f_0(E)$ составляет $(1 + e)^{-1} = 0,27$ и $(1 + 1/e)^{-1} = 0,73$. При $E - F = \pm 2kT$ значения f_0 равны 0,118 и 0,882, а при $E - F = \pm 3kT$ — 0,047 и 0,953. Из этих данных следует, что вероятность заполнения состояний заметно отличается от единицы или нуля лишь в пределах $(2 \div 3) kT$ вблизи значения $E = F$ (рис. 4.2).

Функция распределения Ферми—Дирака характеризует вероятность заполнения данного квантового состояния электроном. Вероятность того, что при тепловом равновесии в состоянии с энергией E электрон отсутствует, т. е. оно занято дыркой, будет равна:

$$f_{0p}(E) = 1 - f_0(E) = 1 - \frac{1}{e^{(E-F)/kT} + 1} = \frac{1}{e^{(F-E)/kT} + 1}. \quad (4.29)$$

Следовательно, функция распределения для дырок аналогична функции распределения для электронов, если отсчитывать энергию

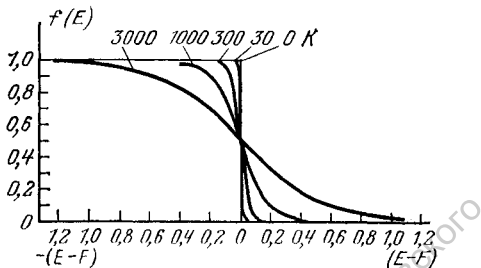


Рис. 4.2. Вид функции распределения Ферми—Дирака

дырок от уровня Ферми в противоположную сторону по сравнению с направлением отсчета энергии для электронов.

Для электронов, находящихся в состояниях с энергией $E - F \gg \gg kT$, выражение для f_0 имеет вид:

$$f_0(E) = e^{(F-E)/kT} = Ce^{-E/kT}, \quad (4.30)$$

т. е. совпадает с функцией распределения Больцмана для частиц, подчиняющихся классической статистике. Если носители заряда подчиняются статистике Больцмана, то электронный газ невырожден и соответственно полупроводник с таким распределением носителей заряда принято называть невырожденным.

4.3. СТЕПЕНЬ ЗАПОЛНЕНИЯ ПРИМЕСНЫХ УРОВНЕЙ

Рассмотрим полупроводник, содержащий донорную примесь в концентрации N_d . Донор, удерживающий электрон, электрически нейтрален. Это соответствует, например, случаю, когда один из узлов кристаллической решетки кремния занят атомом мышьяка. При этом пятый валентный электрон атома донорной примеси не принимает участия в ковалентной связи и ему соответствует энергетический уровень, расположенный ниже дна зоны проводимости на величину E_d (рис. 4.3, а).

Поскольку у донорной примеси имеется только один электрон, который может принимать участие в проводимости, то полное число состояний для донорной примеси должно быть равно количеству атомов введенной примеси на единицу объема кристалла, т. е. равно N_d .

Предположим, что концентрация электронов, находящихся на уровне донорной примеси, равна n_d . В этом случае концентрация ионизованных донорных атомов p_d , образовавшихся в результате тепловых переходов электронов с донорных уровней в зону проводимости и имеющих положительный заряд, составит

$$p_d = N_d - n_d. \quad (4.31)$$

Если бы на примесном донорном уровне согласно принципу Паули могли расположиться два электрона с антипараллельными спинами, то вероятность его заполнения определялась бы функцией Ферми—Дирака (4.28), в которой вместо E следовало поставить

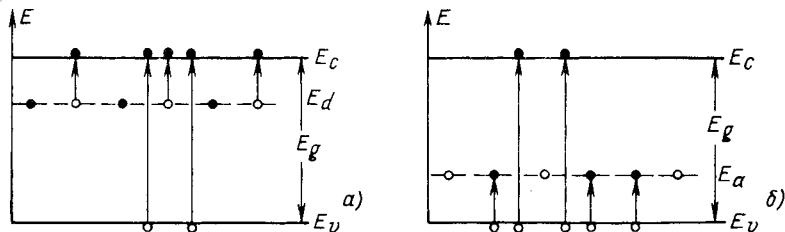


Рис. 4.3. Электронный (а) и акцепторный (б) полупроводники

E_d — энергию электрона на уровне примеси. Но на уровне E_d может быть только один электрон, который может быть захвачен двояким образом в зависимости от направления спина. Следовательно, *нейтральное состояние донорной примеси имеет вдвое больший статистический вес по сравнению с ионизованным состоянием*. Так как при отсутствии электрона на уровне донорной примеси вероятность такого состояния равна 1, то, исходя из принципа Больцмана, можно написать:

$$p_d : n_d = 1 : 2e^{-(E_d - F)/kT}. \quad (4.32)$$

Используя (4.31), это равенство можно записать в виде

$$\frac{N_d - n_d}{n_d} = (2e^{-(E_d - F)/kT})^{-1}, \quad (4.33)$$

откуда следует, что концентрация электронов, находящихся на уровнях донорной примеси, равна:

$$n_d = \frac{N_d}{\frac{1}{2}e^{(E_d - F)/kT} + 1}, \quad (4.34)$$

а концентрация положительных ионов донорной примеси на основании равенств (4.32) и (4.34) будет выражаться соотношением вида

$$p_d = \frac{N_d}{2e^{(F - E_d)/kT} + 1}. \quad (4.35)$$

Предэкспоненциальный множитель в (4.34) в общем случае можно записать через g^{-1} . Тогда вероятность нахождения электрона на донорном уровне с энергией E_d будет определяться выражением

$$f_0(E_d) = \frac{1}{g^{-1}e^{(E_d - F)/kT} + 1}, \quad (4.36)$$

а функция распределения для положительных ионов донорной примеси на основании (4.29) и (4.35) будет:

$$f_{0p} = \frac{1}{ge^{(F - E_d)/kT} + 1}. \quad (4.37)$$

Из (4.36) следует, что для *одновалентной донорной примеси, для которой примесный уровень двукратно вырожден, фактор (степень) спинового вырождения $g = 2$* .

Рассмотрим теперь акцепторный полупроводник, например кремний, легированный бором. Допустим, что концентрация введенной примеси равна N_a . Энергетическая схема такого полупроводника представлена на рис. 4.3, б.

Нейтральный атом бора с соседними атомами кремния образует три ковалентные связи, четвертая связь одного из четырех соседних атомов кремния остается незавершенной, и она, располагаясь около атома бора, ведет себя как положительная дырка. В эту незавершенную связь может перейти электрон от соседнего атома кремния, и для этого потребуется энергия, равная E_a . В результате образуется

свободная дырка, а атом бора превращается в отрицательно заряженный ион бора. Таким образом, на энергетическом уровне акцепторной примеси находится один электрон с произвольным направлением спина (нейтральное состояние акцепторной примеси) либо имеется два электрона с антипараллельными спинами в случае, когда атом акцепторной примеси для укомплектования парной связи захватывает электрон из валентной зоны (ионизованное состояние акцепторной примеси). Следовательно, степень вырождения акцепторного уровня $g = 2$. Поэтому концентрация электронов n_a на уровнях акцепторной примеси (или концентрация отрицательных ионов) при данной температуре будет определяться соотношением вида

$$n_a = \frac{N_a}{g e^{(E_a - F)/kT} + 1}, \quad (4.38)$$

а концентрация дырок на акцепторной примеси p_a соответственно будет равна:

$$p_a = \frac{N_a}{g^{-1} e^{(F - E_a)/kT} + 1}. \quad (4.39)$$

4.4. КОНЦЕНТРАЦИИ ЭЛЕКТРОНОВ И ДЫРОК В ЗОНАХ

Рассмотрим процесс образования свободных носителей заряда под воздействием теплоты в примесном полупроводнике, например в донорном, энергетическая схема которого изображена на рис. 4.3, а.

При температуре абсолютного нуля электроны занимают состояния с минимальным значением энергии. Другими словами, при $T = 0$ К все состояния донорной примеси и в валентной зоне заняты электронами, поэтому в зоне проводимости нет электронов.

При повышении температуры будет происходить отрыв электронов от донорной примеси и разрыв ковалентных связей основного вещества, вследствие чего появятся свободные электроны и дырки. *Свободные носители заряда, возникающие в результате теплового возбуждения и находящиеся с решеткой в термодинамическом равновесии, называют равновесными или тепловыми, а процесс их образования носит название тепловой генерации (или ионизации).*

Следовательно, в донорном полупроводнике при тепловом возбуждении имеет место переход электронов в зону проводимости с донорных уровней и из валентной зоны, где образуются равновесные дырки. Эти переходы на рис. 4.3, а показаны стрелками. *Обратный процесс перехода электрона из зоны проводимости на уровень донорной примеси или в валентную зону называют рекомбинацией.*

В случае акцепторного полупроводника равновесные электроны и дырки появляются в результате теплового перехода электронов из валентной зоны на уровень акцепторной примеси и в зону проводимости (рис. 4.3, б).

Определим при термодинамическом равновесии для примесного полупроводника значения концентрации электронов в зоне прово-

димости n_0 и дырок в валентной зоне p_0 . Решение задачи будем проводить для полупроводника, имеющего в обеих зонах сферические поверхности равной энергии. Равновесную концентрацию электронов можно найти, если в (4.3) подставить плотность квантовых состояний для случая изоэнергетических поверхностей (4.9) и воспользоваться функцией распределения Ферми—Дирака (4.28). Интегрирование нужно провести от дна зоны проводимости E_c до ее потолка. Если учесть резкую зависимость функции $f_0(E)$ от энергии, то верхний предел интегрирования можно положить равным ∞ . В результате будем иметь:

$$n_0 = \int_{E_c}^{\infty} f_0(E) N(E) dE = \int_{E_c}^{\infty} \frac{4\pi (2m_n^*/h^2)^{3/2} (E - E_c)^{1/2} dE}{e^{\frac{E-F}{kT}} + 1}. \quad (4.40)$$

Введем безразмерные величины

$$\frac{E - E_c}{kT} = \varepsilon; \quad \frac{F - E_c}{kT} = \eta. \quad (4.41)$$

Величину η называют приведенным уровнем Ферми. При этом (4.40) преобразуется к виду

$$n_0 = 4\pi \left(\frac{2m_n^*}{h^2} \right)^{3/2} (kT)^{3/2} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} = N_c F_{1/2}(\eta). \quad (4.42)$$

Здесь

$$N_c = 2 \left(\frac{2\pi m_n^* kT}{h^2} \right)^{3/2} \quad (4.43)$$

— эффективная плотность состояний в зоне проводимости;

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} \quad (4.44)$$

— интеграл Ферми—Дирака с индексом $1/2$. Он представлен на рис. 4.4.

Эффективная плотность состояний N_c зависит от температуры. Если в (4.43) подставить численные значения универсальных констант, то будем иметь:

$$\begin{aligned} N_c &= 2 (2\pi m_n^* kT/h^2)^{3/2} = 4,82 \cdot 10^{15} (m_n^*/m_0)^{3/2} T^{3/2} = \\ &= 2,5 \cdot 10^{19} (m_n^*/m_0)^{3/2} (T/300)^{3/2}, \end{aligned} \quad (4.45)$$

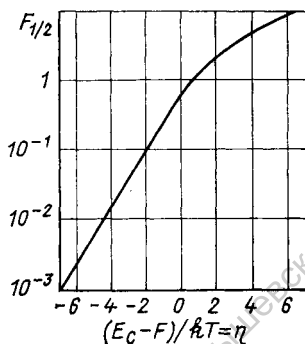


Рис. 4.4. Зависимость интеграла Ферми—Дирака от приведенного уровня Ферми

Определим концентрацию дырок в валентной зоне. В соответствии с (4.11) и (4.29) получаем:

$$p_0 = \int_{E_{vmin}}^{E_v} f_{0p}(E) \Lambda(E) dE = 4\pi \left(\frac{2m_p^*}{h^2} \right)^{3/2} \int_{-\infty}^{E_v} \frac{(E_v - E)^{1/2} dE}{e^{(F-E)/kT} + 1}. \quad (4.46)$$

В этом равенстве нижний предел интегрирования E_{vmin} заменили на $-\infty$, учитывая резкую зависимость f_{0p} от энергии.

Обозначим:

$$\frac{E_v - E}{kT} = \varepsilon_p; \quad \frac{E_c - E_v}{kT} = \frac{E_g}{kT} = \varepsilon_i; \quad \frac{F - E}{kT} = \eta + \varepsilon_i + \varepsilon_p. \quad (4.47)$$

Здесь $E_c - E_v = E_g$ — ширина запрещенной зоны. Теперь выражение для концентрации дырок примет вид:

$$p_0 = \frac{4}{\sqrt{\pi}} \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2} \int_0^\infty \frac{\varepsilon_p^{1/2} d\varepsilon_p}{e^{\varepsilon_p + \eta + \varepsilon_i} + 1} = N_v F_{1/2}(-\eta - \varepsilon_i), \quad (4.48)$$

где

$$N_v = 2 \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2} \quad (4.49)$$

— эффективная плотность состояний в валентной зоне;

$$F_{1/2}(-\eta - \varepsilon_i) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\varepsilon_p^{1/2} d\varepsilon_p}{e^{\varepsilon_p + \eta + \varepsilon_i} + 1} \quad (4.50)$$

— интеграл Ферми — Дирака для валентной зоны.

Если положить $m_n^* = m_p^* = m_0$, то при $T = 300$ К $N_c = N_v = 2,5 \cdot 10^{19}$ см⁻³. Это составляет приблизительно $1/2000$ числа атомов в 1 см³.

Если ввести обозначение $\frac{E_c - E_d}{kT} = \varepsilon_d$, то на основании (4.36) концентрация электронов на уровне донорной примеси будет равна:

$$n_d = f_0(E_d) N_d = \frac{N_d}{g^{-1} e^{-\varepsilon_d - \eta} + 1}, \quad (4.51)$$

а количество положительных ионов донорной примеси согласно (4.37) составит:

$$p_d = f_{0p}(E_d) N_d = \frac{N_d}{g e^{\varepsilon_d + \eta} + 1}. \quad (4.52)$$

Определим концентрации дырок p_a и электронов n_a на акцепторной примеси. Если обозначить $\frac{E_c - E_a}{kT} = \epsilon_a$, то на основании (4.38) и (4.39)

$$p_a = \frac{N_a}{g^{-1} e^{\epsilon_a + \eta} + 1}; \quad (4.53)$$

$$n_a = \frac{N_a}{ge^{-\epsilon_a - \eta} + 1}. \quad (4.54)$$

4.5. ПРИМЕСНЫЙ ПОЛУПРОВОДНИК

Для того чтобы определить равновесные концентрации электронов n_0 и дырок p_0 в зонах, необходимо вычислить интегралы Ферми—Дирака (4.44) и (4.50), которые в зависимости от приведенного уровня Ферми можно выразить следующим образом.

1. При $-\infty < \eta < -1$

$$F_{1/2}(\eta) = e^\eta, \quad (4.55)$$

а при $-\infty < (-\eta - \epsilon_i) < -1$

$$F_{1/2}(-\eta - \epsilon_i) = e^{-\eta - \epsilon_i}. \quad (4.56)$$

Такие приближения соответствуют статистике Больцмана, а полупроводник является невырожденным. Это классическое приближение, для которого равновесные концентрации электронов и дырок соответственно равны:

$$n_0 = N_c e^\eta; \quad (4.57)$$

$$p_0 = N_v e^{-\eta - \epsilon_i}. \quad (4.58)$$

На рис. 4.5 классическое приближение изображено кривой 1.

2. При $5 < \eta < \infty$

$$F_{1/2}(\eta) = \frac{4}{3\sqrt{\pi}} \eta^{3/2}, \quad (4.59)$$

а при $5 < (-\eta - \epsilon_i) < \infty$

$$F_{1/2}(-\eta - \epsilon_i) = \frac{4}{3\sqrt{\pi}} (-\eta - \epsilon_i)^{3/2}. \quad (4.60)$$

Эти приближения справедливы для случая сильного вырождения донорного и акцепторного полупроводников соответственно. На рис. 4.5 приближение сильного вырождения донорного полупроводника дано кривой 2.

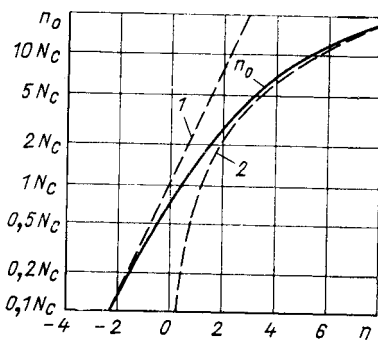


Рис. 4.5. Зависимость концентрации электронов проводимости от положения приведенного уровня Ферми

3. При $-1 < \eta < 5$ для зоны проводимости можно приближенно записать

$$F_{1/2}(\eta) \approx \frac{1}{0,17 + e^\eta}, \quad (4.61)$$

а при $-1 < (-\eta - \varepsilon_i) < 5$ для валентной зоны

$$F_{1/2}(-\eta - \varepsilon_i) = \frac{1}{0,27 + e^{-\eta - \varepsilon_i}}. \quad (4.62)$$

При этих условиях свойства полупроводника являются переходными от невырожденного к вырожденному. Расчет интегралов Ферми—Дирака по формулам (4.61) и (4.62) дает значительную погрешность, а поэтому целесообразно пользоваться специальными таблицами интегралов Ферми—Дирака.

На рис. 4.5 кривая n_0 представляет собой истинное значение концентрации равновесных электронов проводимости, подсчитанное по соотношению (4.42).

Рассмотрим более подробно крайние случаи.

Невырожденный примесный полупроводник. Согласно рис. 4.5 можно считать, что донорный полупроводник невырожден, если

$$\eta = \frac{F - E_c}{kT} < -1$$

или

$$F < (E_c - kT). \quad (4.63)$$

Это означает, что в невырожденном донорном полупроводнике уровень Ферми лежит в запрещенной зоне ниже дна зоны проводимости по крайней мере на kT . В этом случае для электронов, находящихся в зоне проводимости и обладающих энергией $E \geq E_c$, в функции распределения Ферми—Дирака экспоненциальный член будет намного больше единицы, а поэтому ею можно пренебречь, т. е.

$$f_0 = \frac{1}{e^{(E-F)/kT} + 1} \approx e^{-(E-F)/kT}. \quad (4.64)$$

Из соотношения (4.64) следует, что в невырожденном донорном полупроводнике носители заряда подчиняются статистике Больцмана. В соответствии с этим интеграл Ферми—Дирака запишется в виде

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} = \frac{2}{\sqrt{\pi}} e^\eta \int_0^\infty e^{-\varepsilon} \varepsilon^{1/2} d\varepsilon = e^\eta = e^{-(E_c - F)/kT}. \quad (4.65)$$

Равновесные концентрации электронов n_0 и дырок p_0 для невырожденного полупроводника соответственно равны:

$$n_0 = N_c F_{1/2}(\eta) = N_c e^\eta = N_c e^{-(E_c - F)/kT}; \quad (4.66)$$

$$p_0 = N_v F_{1/2}(-\eta - \varepsilon_i) = N_v e^{-\eta - \varepsilon_i} = N_v e^{(E_v - F)/kT}. \quad (4.67)$$

Анализ этих соотношений показывает, что для невырожденного распределения носителей заряда значение интеграла Ферми совпадает с выражением вероятности заполнения уровня энергии E_c . Следовательно, как вытекает из (4.66) и (4.67), в невырожденном примесном полупроводнике концентрация электронов в зоне проводимости такая, какая была бы на N_c уровнях с энергией E_c , а количество свободных дырок такое же, как если бы вместо валентной зоны было N_v уровней и все с энергией E_v .

Акцепторный полупроводник невырожден при $(-\eta - \epsilon_i) < -1$ или $F > (E_v + kT)$. Следовательно, дырочный полупроводник невырожден, если уровень Ферми лежит в запрещенной зоне выше потолка валентной зоны не менее чем на kT , вследствие чего равновесные концентрации электронов и дырок подчиняются статистике Больцмана и определяются равенствами (4.66) и (4.67).

В собственном полупроводнике свободные носители заряда возникают в результате разрыва ковалентных связей, поэтому концентрации свободных электронов и дырок равны: $n_0 = p_0 = n_i$. Величина n_i называется собственной концентрацией носителей заряда. Если для собственного полупроводника обозначить приведенный уровень Ферми через η_i , то на основании (4.66) и (4.67) следует:

$$n_i = N_c e^{\eta_i} = N_v e^{-\eta_i - \epsilon_i}. \quad (4.68)$$

Выразим теперь равновесные концентрации электронов n_0 (4.66) и дырок p_0 (4.67) для невырожденного примесного полупроводника через n_i . Тогда

$$\left. \begin{aligned} n_0 &= N_c e^{\eta} = n_i e^{\eta - \eta_i}; \\ p_0 &= N_v e^{-\eta - \epsilon_i} = n_i e^{\eta_i - \eta} \end{aligned} \right\}. \quad (4.69)$$

Отсюда следует:

$$n_0 p_0 = n_i^2. \quad (4.70)$$

Таким образом, в невырожденном примесном полупроводнике произведение концентраций свободных электронов и дырок при термодинамическом равновесии есть постоянная величина, равная квадрату концентрации носителей заряда в собственном полупроводнике при данной температуре.

Зонная диаграмма, плотность состояний $N(E)$, распределение Ферми—Дирака $f_0(E)$ и концентрация носителей заряда в собственном, электронном и дырочном полупроводниках при термодинамическом равновесии показаны на рис. 4.6.

Если носители заряда появляются в результате теплового возбуждения, то вероятность такого процесса пропорциональна $e^{-\omega/kT}$. Величину ω называют энергией активации. Поэтому концентрацию электронов n_0 можно выразить и через энергию активации. Она будет равна:

$$n_0 = C e^{-\omega/kT}, \quad (4.71)$$

где C — постоянная величина.

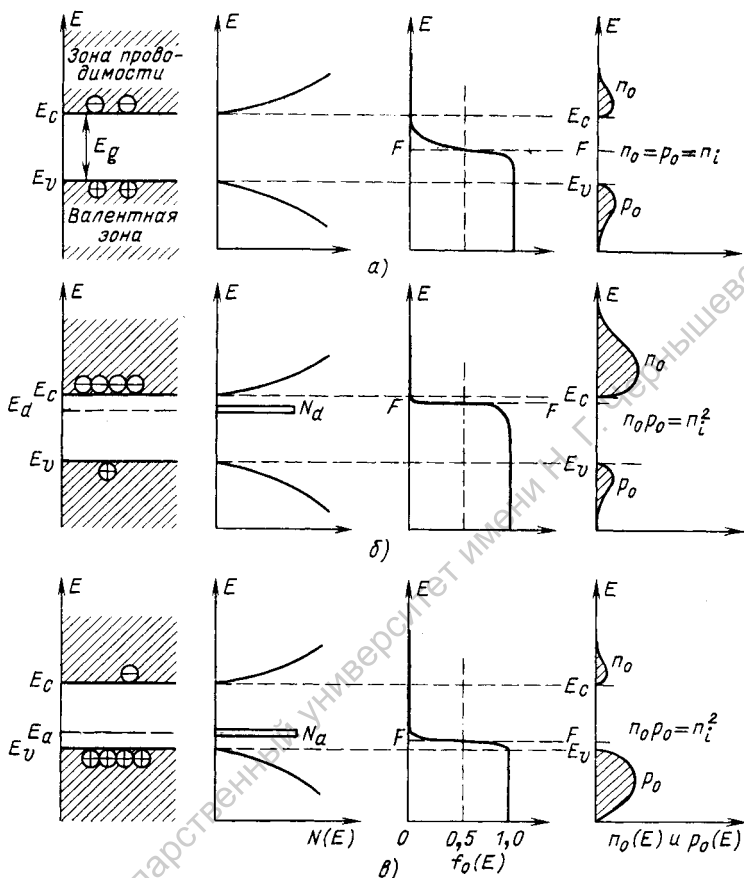


Рис. 4.6. Зонная диаграмма, плотность состояний, распределение Ферми—Дирака и концентрация носителей заряда в собственном (а), донорном (б) и акцепторном (в) полупроводниках при термодинамическом равновесии

Соотношение (4.66) может быть записано как

$$n_0 T^{-3/2} = 2 \left(\frac{2\pi m_n^* k}{h^2} \right)^{3/2} e^{-(E_c - F)/kT} = C e^{-(E_c - F)/kT}.$$

Из сравнения этой формулы с (4.71) следует, что

$$w = -k \frac{d}{d\left(\frac{1}{T}\right)} \ln(n_0 T^{-3/2}) = E_c - F + T \frac{dF}{dT}, \quad (4.72)$$

т. е. энергия активации зависит от положения уровня Ферми.

Сильно вырожденный примесный полупроводник. Согласно (4.59) можно считать, что донорный полупровод-

ник сильно вырожден при $\eta = \frac{F - F_c}{kT} > 5$, т. е. когда уровень Ферми расположен в зоне проводимости выше дна зоны проводимости E_c по крайней мере на $5kT$.

Если в зоне проводимости при $T = 0$ К имеются свободные электроны, то полупроводник, для которого приведенный уровень Ферми $\eta = \frac{F - F_c}{kT} \rightarrow \infty$, будет сильно вырожденным. В этом случае в выражении (4.42) знаменатель равен единице и равновесная концентрация электронов в сильно вырожденном донорном полупроводнике будет равна:

$$\begin{aligned} n_0 &= N_c \frac{2}{\sqrt{\pi}} \int_0^{\eta_0} \varepsilon^{1/2} d\varepsilon = \frac{4}{3\sqrt{\pi}} N_c \eta_0^{3/2} = \frac{4}{3\sqrt{\pi}} N_c \left(\frac{F_0 - E_c}{kT} \right)^{3/2} = \\ &= \frac{8\pi}{3} \left(\frac{2m_n^*}{\hbar^2} \right)^{3/2} (F_0 - E_c)^{3/2}, \end{aligned} \quad (4.73)$$

а концентрация равновесных дырок p_0 определяется (4.67).

Уровень Ферми при температуре абсолютного нуля, как следует из (4.73) имеет вид

$$F_0 = E_c + \frac{\hbar^2}{2m_n^*} \left(\frac{3n_0}{8\pi} \right)^{2/3}. \quad (4.74)$$

Акцепторный полупроводник является сильно вырожденным, если $(-\eta - \varepsilon_i) > 5$ и $F < (E_v - 5kT)$, т. е. если уровень Ферми лежит в валентной зоне ниже E_v на $5kT$. В сильно вырожденном акцепторном полупроводнике равновесная концентрация свободных дырок в соответствии с (4.48) и (4.60) равна:

$$p_0 = N_v \frac{4}{3\sqrt{\pi}} (-\eta_0 - \varepsilon_i)^{3/2} = \frac{8\pi}{3} \left(\frac{2m_p^*}{\hbar^2} \right)^{3/2} (E_v - F_0)^{3/2}, \quad (4.75)$$

а концентрация электронов проводимости n_0 определяется соотношением (4.66).

Из (4.73) и (4.75) следует, что концентрация основных носителей заряда в сильно вырожденном примесном полупроводнике не зависит от температуры. Положение уровня Ферми F_0 определяется концентрацией равновесных носителей заряда и зависит от их эффективной массы. При этом, чем меньше эффективная масса носителей заряда, тем соответственно и меньше их концентрация, при которой наступает вырождение.

Условие полного вырождения с учетом (4.74) можно записать следующим образом:

$$\eta = \frac{F - E_c}{kT} = \frac{\hbar^2}{2m_n^* kT} \left(\frac{3n_0}{8\pi} \right)^{2/3} \gg 1. \quad (4.76)$$

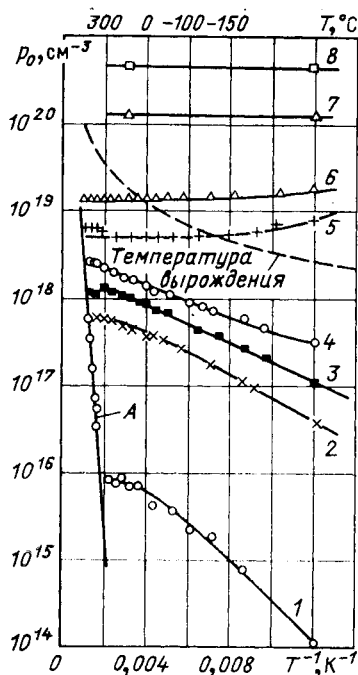


Рис. 4.7. Температурная зависимость концентрации дырок для кремния, легированного бором:

А — собственная концентрация. Концентрация атомов бора для различных кривых равна: 1 — нелегированный; 2 — $6,7 \cdot 10^{17}$; 3 — $1,3 \times 10^{18}$; 4 — $2,7 \cdot 10^{18}$; 5 — $6,7 \cdot 10^{18}$; 6 — $1,3 \times 10^{19}$; 7 — $1,3 \cdot 10^{20}$; 8 — $1,3 \cdot 10^{21}$ см $^{-3}$ (пунктирная кривая показывает связь концентрации дырок с температурой, при которой наступает вырождение)

Здесь η_0 заменена η , что, как будет показано ниже, для случая сильного вырождения не вносит существенной погрешности. Из анализа (4.76) следует, что вырождение наступает тем раньше, чем больше концентрация свободных носителей заряда, чем ниже температура (при данной концентрации) и чем меньше эффективная масса носителей заряда.

Граница перехода от невырожденного полупроводника к вырожденному определяется концентрацией носителей заряда вырождения n_v при температуре вырождения T_v . Поскольку n_v

является минимальной концентрацией носителей заряда, соответствующей началу вырождения при данной температуре, т. е. при $F_0 = E_c + kT$, на основании (4.73) можно записать:

$$n_v = 4N_c/3 \sqrt{\pi} = 3,78 \cdot 10^{15} (m_n^*/m_0)^{3/2} T_v^{3/2} = 2 \cdot 10^{19} (m_n^*/m_0)^{3/2} (T_v/300)^{3/2} \quad (4.77)$$

или

$$T_v = 4,1 \cdot 10^{-11} (m_0/m_n^*) n_v^{2/3}. \quad (4.78)$$

Эти выводы подтверждаются экспериментальными данными. На рис. 4.7 представлены кривые температурной зависимости концентрации дырок для кремния с разным содержанием бора.

Анализ этих зависимостей показывает, что с увеличением степени легирования концентрация дырок перестает зависеть от температуры, т. е. полупроводник становится сильно вырожденным. Область вырождения с увеличением температуры сдвигается к более высоким концентрациям носителей заряда. Если при температуре жидкого азота вырождение наступает при $p_0 = 2,5 \cdot 10^{18}$ см $^{-3}$, то при комнатной температуре оно имеет место при $p_0 = 2 \cdot 10^{19}$ см $^{-3}$. Антимонид индия, обладающий очень малой эффективной массой электронов ($m_n^* = 0,013m_0$), при $T = 300$ К оказывается вырожденным уже при $n_0 = 2 \cdot 10^{18}$ см $^{-3}$.

Определим зависимость уровня Ферми для вырожденного донорного полупроводника от температуры. Если $\eta \gg 1$, то $e^{e-\eta} \ll 1$ и интеграл Ферми можно разложить в степенной ряд, а поэтому равновесная концентрация электронов n_0 может быть представлена формулой

$$n_0 = N_c \frac{4}{3\sqrt{\pi}} \eta^{3/2} \left(1 + \frac{\pi^2}{8\eta^2} - \frac{7\pi^4}{640\eta^4} + \frac{31\pi^6}{3072\eta^6} - \dots \right). \quad (4.79)$$

Отбросим члены ряда, начиная с третьего, и заменим во втором слагаемом η на η_0 . Можно показать, что при этом несколько снижается погрешность, возникающая при отбрасывании последних членов. В результате для n_0 получим приближенное выражение

$$n_0 \approx N_c \frac{4}{3\sqrt{\pi}} \eta^{3/2} \left(1 + \frac{\pi^2}{8\eta_0^2} \right). \quad (4.80)$$

Из соотношений (4.73) и (4.80) имеем:

$$\frac{3\sqrt{\pi}}{4} \frac{n_0}{N_c} = \eta^{3/2} \left(1 + \frac{\pi^2}{8\eta_0^2} \right) = \eta_0^{3/2}, \quad (4.81)$$

откуда

$$\eta = \eta_0 \left(1 + \frac{\pi^2}{8\eta_0^2} \right)^{-2/3}. \quad (4.81a)$$

При $\eta_0 > 5$ второй член в скобке мал. Разложив (4.81a) в ряд и ограничившись двумя членами, будем иметь:

$$\eta = \eta_0 \left(1 - \frac{\pi^2}{12\eta_0^2} \right) \quad (4.82)$$

или

$$F = F_0 \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{F_0 - E_c} \right)^2 \right]. \quad (4.83)$$

Как следует из (4.83), при достаточно низкой температуре уровень Ферми F в вырожденном полупроводнике очень близок к F_0 .

4.6. СОБСТВЕННЫЙ ПОЛУПРОВОДНИК

Полупроводник будет являться собственным, если влияние примесей для него пренебрежимо мало. В нем свободные носители заряда возникают только за счет разрыва валентных связей (рис. 4.8), поэтому

$$n_0 = p_0 \quad (4.84)$$

Соотношение (4.84) выражает собой условие электронейтральности собственного полупроводника, по которому суммарный заряд всех заряженных частиц кристалла должен быть равен нулю. Это требование определяет положение уровня Ферми. Его можно найти, решив уравнение

$$N_c F_{1/2}(\eta) = N_v F_{1/2}(-\eta - \epsilon_i). \quad (4.85)$$

Невырожденный собственный полупроводник. Для него зонная диаграмма, плотность состояний, распреде-

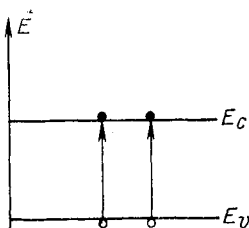


Рис. 4.8. Тепловая генерация носителей заряда в собственном полупроводнике

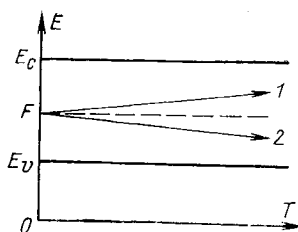


Рис. 4.9. Изменение уровня Ферми с температурой в собственном полупроводнике

ление Ферми—Дирака и концентрация носителей заряда при термодинамическом равновесии представлены на рис. 4.6, а. Для невырожденного полупроводника интеграл Ферми—Дирака можно аппроксимировать экспонентой (4.65). При этом условие (4.85) примет вид:

$$N_c e^{\eta} = N_v e^{-\eta - \epsilon_i}, \quad (4.86)$$

откуда легко получить для уровня Ферми

$$F = E_c - \frac{1}{2} E_g + \frac{kT}{2} \ln \frac{N_v}{N_c} = \frac{E_c + E_v}{2} + \frac{kT}{2} \ln \frac{N_v}{N_c}, \quad (4.87)$$

где $E_g = E_c - E_v$ — ширина запрещенной зоны.

Если подставить в (4.87) значения N_c из (4.43) и N_v из (4.49), то получим:

$$F = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln \frac{m_p^*}{m_n^*}. \quad (4.88)$$

Из (4.88) следует, что при температуре абсолютного нуля уровень Ферми для собственного полупроводника располагается посередине между дном зоны проводимости и потолком валентной зоны, т. е. $F = \frac{1}{2} (E_c + E_v)$. В случае $m_p^* = m_n^*$ его положение не зависит от температуры и F лежит в середине запрещенной зоны. При $m_p^* > m_n^*$ уровень Ферми с повышением температуры линейно смещается к дну зоны проводимости, а при $m_p^* < m_n^*$ — к потолку валентной зоны (рис. 4.9, кривые 1 и 2 соответственно).

Определим собственную концентрацию носителей заряда:

$$\begin{aligned} n_i &= (n_0 p_0)^{1/2} = (N_c N_v)^{1/2} e^{-(E_c - E_v)/2kT} = \\ &= 2 \left(\frac{2\pi k}{h^2} \right)^{3/2} (m_n^* m_p^*)^{3/4} T^{3/2} e^{-E_g/2kT} = 4,9 \cdot 10^{15} \left(\frac{m_n^* m_p^*}{m_0^2} \right)^{3/4} T^{3/2} e^{-E_g/2kT}. \end{aligned} \quad (4.89)$$

Таким образом, концентрация носителей заряда в невырожденном собственном полупроводнике зависит от температуры T , ширины

запрещенной зоны E_g , значений эффективных масс электронов m_n^* и дырок m_p^* и не зависит от положения уровня Ферми F . Температурная зависимость собственной концентрации носителей заряда при $E_g \gg kT$ определяется в основном экспоненциальным членом (4.89). Доказательством этого являются данные рис. 4.10, на котором представлена зависимость собственной концентрации носителей заряда в германии, кремнии и арсениде галлия от обратной температуры.

Энергия активации для собственного полупроводника, как следует из (3.72) и (3.87),

$$\omega = \frac{1}{2} E_g, \quad (4.90)$$

т. е. равна половине ширины запрещенной зоны.

Оценим собственную концентрацию носителей заряда в кремнии, германии и антимониде индия. Для кремния $m_{dn}^* = 1,08m_0$, $m_{dp}^* = 0,56m_0$ при $T = 300$ К $E_g = 1,11$ эВ, поэтому $n_i = 1,9 \cdot 10^{10}$ см⁻³. Для германия $m_{dn}^* = 0,56m_0$, $m_{dp}^* = 0,35m_0$, $E_g = 0,67$ эВ и $n_i = 2 \cdot 10^{13}$ см⁻³. Для антимонида индия $m_n^* = 0,013m_0$, $m_p^* = 0,6m_0$, $E_g = 0,18$ эВ и $n_i = 2 \cdot 10^{16}$ см⁻³.

Согласно (4.89) графическая зависимость $\ln(n_i T^{-3/2})$ от $1/T$ должна выражаться прямой линией:

$$\ln(n_i T^{-3/2}) = \text{const} - \frac{F_g}{2k} \frac{1}{T}. \quad (4.91)$$

Угол наклона этой прямой $\text{tg } \theta = E_g/2k$. Однако полученное по данным рис. 4.10 значение E_g не соответствует истинной ширине запрещенной зоны германия. Причина этого расхождения заключается в том, что сама ширина запрещенной зоны изменяется с температурой. Рост амплитуды тепловых колебаний атомов решетки приводит к ее уменьшению. Кроме того, с увеличением температуры изменяются межатомные расстояния, что также оказывает влияние на ширину запрещенной зоны. В результате зависимость $E_g(T)$ может иметь сложный характер. В качестве примера на рис. 4.11 показаны изменения ширины запрещенной зоны в зависимости от температуры для германия, кремния и арсенида галлия. Для этих полупроводников значения ширины запрещенной зоны при 0 К составляют 0,89; 1,16 и 1,52 эВ соответственно. У них, как следует из

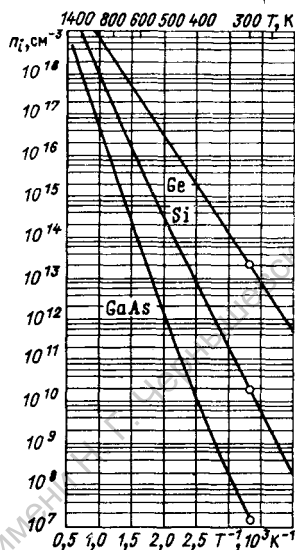


Рис. 4.10. Зависимость собственной концентрации носителей заряда в германии, кремнии и арсениде галлия от обратной температуры

рис. 4.11, в диапазоне температур 175—350 К ширина запрещенной зоны меняется линейно с температурой. При этом температурный коэффициент изменения ширины запрещенной зоны $\alpha = dE_g/dT < 0$ и равен $3,9 \cdot 10^{-4}$; $2,4 \cdot 10^{-4}$; $2,9 \cdot 10^{-4}$ и $4,3 \cdot 10^{-4}$ эВ/К соответственно для германия, кремния, антимоноида индия и арсенида галлия. У PbS $\alpha > 0$ возрастает от 0,34 эВ при 0 К до 0,41 эВ при 300 К. В этих случаях зависимость подчиняется линейному закону вида

$$E_g = E_{g0} - \alpha T, \quad (4.92)$$

где E_{g0} — экстраполированная ширина запрещенной зоны при 0 К (рис. 4.11, например, для GaAs $E_{g0} = 1,58$ эВ).

Подставив (4.92) в уравнение (4.89), получим:

$$n_i T^{-3/2} = 2 \left(\frac{2\pi^2 m_n^* m_p^* k^2}{h^2} \right)^{3/4} e^{-\alpha/2k} e^{-E_{g0}/2kT}, \quad (4.93)$$

откуда найдем, что

$$\ln(n_i T^{-3/2}) = \text{const} - \frac{\alpha}{2k} - \frac{E_{g0}}{2k} \frac{1}{T}. \quad (4.94)$$

Таким образом, если ширина запрещенной зоны полупроводника линейно зависит от температуры, график зависимости $\ln(n_i T^{-3/2})$ от $1/T$ также представляет собой прямую линию, наклон которой характеризуется значением E_{g0} , которое является экстраполированной шириной запрещенной зоны при 0 К. Истинное значение ширины запрещенной зоны полупроводника при данной температуре определяется по формуле (4.92).

В вырожденных полупроводниках. У собственного полупроводника согласно (4.88) скорость изменения уровня Ферми с температурой пропорциональна отношению эффективных масс дырок и электронов. В результате этого с повышением температуры уровень Ферми отдалается от зоны с тяжелыми носителями заряда, приближаясь к зоне с легкими носителями заряда. И если расстояние от уровня Ферми до этой зоны становится соизмеримо со значением kT , то

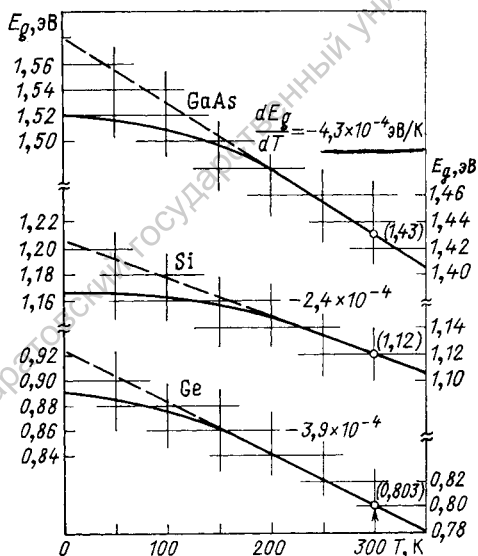


Рис. 4.11. Зависимость ширины запрещенной зоны германия, кремния и арсенида галлия от температуры

в ней наступает вырождение и соответствующий интеграл Ферми — Дирака уже не может быть заменен экспонентой. При этом чем сильнее различаются эффективные массы электронов и дырок, тем раньше наступает вырождение. Если, например, вырождение наступило в зоне проводимости, то в валентной зоне оно отсутствует и выражение для собственной концентрации носителей будет иметь вид:

$$n_i = N_c F_{1/2}(\eta) = N_c e^{-\eta} e^{-\varepsilon_i}. \quad (4.95)$$

Отсюда получается следующая формула, определяющая положение уровня Ферми:

$$e^{\eta} F_{1/2}(\eta) = (m_p^*/m_n^*)^{3/2} e^{-\varepsilon_i}. \quad (4.96)$$

Классическим примером полупроводника с сильно различающимися эффективными массами является антимоид индия, для которого $m_p^* \approx 10 m_n^*$. На рис. 4.12 показано изменение положения приведенного уровня Ферми в InSb в зависимости от температуры. При этом учитывается зависимость $E_g(T)$, которая согласно экспериментальным данным имеет вид:

$$E_g = (0,26 - 2,7 \cdot 10^{-4} T) \text{ эВ}.$$

Как следует из рис. 4.12, уже при $T = 200 \text{ К}$ уровень Ферми лежит ниже E_c всего на $4 kT$, а около 450 К пересекает уровень E_c и входит в зону проводимости, что приводит к сильному вырождению электронного газа. Это необходимо учитывать, и при расчете n_i следует использовать равенство (4.95).

4.7. ЗАВИСИМОСТЬ УРОВНЯ ФЕРМИ ОТ КОНЦЕНТРАЦИИ ПРИМЕСИ И ТЕМПЕРАТУРЫ ДЛЯ НЕВЫРОЖДЕННОГО ПОЛУПРОВОДНИКА

Полупроводник может содержать как донорную, так и акцепторную примеси. Пусть концентрации этих примесей соответственно

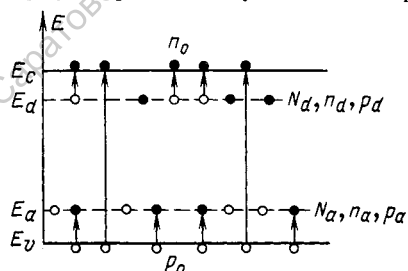


Рис. 4.13. Полупроводник, имеющий примеси донорного и акцепторного типа

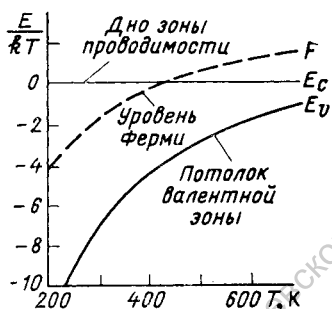


Рис. 4.12. Положение уровня Ферми (в единицах kT) в зависимости от температуры для антимоида индия

равны N_d и N_a . Энергетическая схема такого полупроводника представлена на рис. 4.13. Электроны в зоне проводимости и дырки в валентной зоне возникают в результате теплового возбуждения электронов из валентной зоны и атомов примеси. Эти переходы на рисунке показаны стрелками.

Для определения положения уровня Ферми в полупроводнике

при термодинамическом равновесии можно использовать условие электронейтральности, по которому в любой точке кристалла суммарный заряд всех заряженных частиц должен быть равен нулю. Для полупроводника с двумя видами примеси уравнение электронейтральности запишется в виде

$$n_0 + n_a = p_0 + p_d, \quad (4.97)$$

где $n_a = N_a - p_a$ — количество электронов, связанных с акцепторной примесью; $p_d = N_d - n_d$ — число ионизованных атомов донорной примеси. С учетом выражений для n_a и p_d условие электронейтральности (4.97) можно представить в виде

$$(n_0 + n_d) - (p_0 + p_a) = N_d - N_a. \quad (4.98)$$

Выразив концентрации носителей заряда в соответствии с формулами (4.42), (4.48), (4.51) и (4.53), получим:

$$N_c F_{1/2}(\eta) + \frac{N_d}{g^{-1} e^{-\varepsilon_d - \eta} + 1} - N_v F_{1/2}(-\eta - \varepsilon_i) - \frac{N_a}{g^{-1} e^{\varepsilon_a + \eta} + 1} = N_d - N_a. \quad (4.99)$$

Для невырожденного полупроводника равенство (4.99) значительно упрощается

$$N_c e^{\eta} + \frac{N_d}{g^{-1} e^{-\varepsilon_d - \eta} + 1} - N_v e^{-\eta - \varepsilon_i} - \frac{N_a}{g^{-1} e^{\varepsilon_a + \eta} + 1} = N_d - N_a, \quad (4.100)$$

однако и в этом случае определение уровня Ферми встречает большие трудности. Поэтому в дальнейшем рассмотрим решение частных задач, имеющих тем не менее большое практическое значение.

1. Донорный полупроводник ($N_a = 0$)

Ограничимся вначале областью температур, при которых имеет место лишь ионизация примесных центров, а собственная проводимость отсутствует, т. е. $p_0 = 0$. Условие электронейтральности (4.98) запишется при этом так:

$$n_0 + n_d = N_d \quad (4.101)$$

или

$$n_0 = p_d,$$

откуда согласно (4.52) будем иметь:

$$n_0 = \frac{N_d}{g e^{\varepsilon_d + \eta} + 1}. \quad (4.102)$$

Считаем, что вырождение отсутствует, следовательно, $n_0 = N_c e^{\eta}$. Из этого равенства найдем, что

$$e^{\eta} = n_0 / N_c. \quad (4.103)$$

Подставив (4.103) в (4.102), получим уравнение для концентрации электронов в зоне проводимости:

$$n_0 = \frac{N_c N_d}{N_c + g e^{\varepsilon_d} n_0}, \quad (4.104)$$

которое легко преобразуется к виду

$$n_0^2 + g^{-1} e^{-\varepsilon_d} N_c n_0 - g^{-1} e^{-\varepsilon_d} N_c N_d = 0. \quad (4.105)$$

Решая это квадратное уравнение относительно n_0 , находим:

$$n_0 = \frac{g^{-1} e^{-\varepsilon_d} N_c}{2} (\sqrt{1 + 4 g e^{\varepsilon_d} N_c^{-1} N_d} - 1). \quad (4.106)$$

Минус перед корнем отброшен, так как подкоренное выражение больше единицы, а $n_0 > 0$. Это соотношение можно свести к следующему выражению:

$$n_0 = \frac{2 N_d}{\sqrt{1 + 4 g e^{\varepsilon_d} N_c^{-1} N_d} + 1}. \quad (4.107)$$

Положение уровня Ферми можно определить из соотношения (4.103)

$$e^{\eta} = \frac{n_0}{N_c} = \frac{2 N_d / N_c}{\sqrt{1 + 4 g e^{\varepsilon_d} N_c^{-1} N_d} + 1}. \quad (4.108)$$

Прологарифмировав равенство (4.108), получим:

$$\eta = -\ln \left[\frac{N_c}{2 N_d} + \sqrt{\left(\frac{N_c}{2 N_d} \right)^2 + g e^{\varepsilon_d} N_c N_d^{-1}} \right] \quad (4.109)$$

или

$$F = E_c - kT \ln \left[\frac{N_c}{2 N_d} + \sqrt{\left(\frac{N_c}{2 N_d} \right)^2 + g e^{\varepsilon_d} N_c N_d^{-1}} \right]. \quad (4.110)$$

Проведем анализ выражения (4.110) для уровня Ферми. В области низких температур, где справедливо условие $e^{(E_c - E_d)/kT} \gg N_c / 2 N_d$, формула (4.110) упрощается и положение уровня Ферми будет определяться соотношением вида

$$F = E_c - \frac{E_c - E_d}{2} - \frac{kT}{2} \ln \frac{g N_c}{N_d} = \frac{E_c + E_d}{2} - \frac{kT}{2} \ln \frac{g N_c}{N_d}, \quad (4.111)$$

а энергия активации согласно (4.72) будет:

$$\omega = \frac{1}{2} (E_c - E_d),$$

т. е. равна половине энергии ионизации донорной примеси.

Как следует из соотношения (4.111), в невырожденном донорном полупроводнике при температуре абсолютного нуля уровень Ферми располагается посередине между дном зоны проводимости и уровнем донорной примеси. В области достаточно низких температур (не-

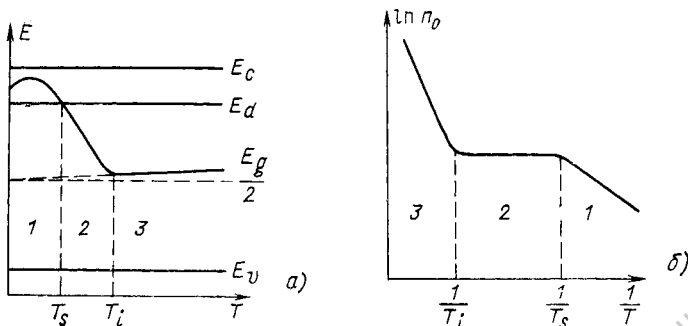


Рис. 4.14. Изменение положения уровня Ферми (а) и концентрации электронов (б) с температурой для донорного полупроводника

сколько градусов по шкале Кельвина), когда $gN_c < N_d$, уровень Ферми вначале повышается до некоторого максимального значения, а затем начинает снижаться и при $gN_c = N_d$ снова имеем $F = \frac{1}{2}(E_c + E_d)$, как и для случая $T = 0$. Дальнейшее повышение температуры сопровождается ростом N_c и в области температуры, когда $gN_c > N_d$, уровень Ферми продолжает снижаться. Такому перемещению уровня Ферми соответствует экспоненциальная температурная зависимость концентрации электронов

$$n_0 = N_c e^n = N_c e^{-(E_c - E_d)/2kT - \frac{1}{2} \ln(gN_c/N_d)} = (g^{-1} N_c N_d)^{1/2} e^{-(E_c - E_d)/2kT}. \quad (4.112)$$

Эта область изменения уровня Ферми с температурой, которая описывается формулой (4.111), является областью слабой ионизации примеси (или областью выморозивания). Она обозначена цифрой 1 на рис. 4.14, на котором проиллюстрировано изменение уровня Ферми и концентрации электронов в зависимости от температуры для донорного полупроводника.

При дальнейшем повышении температуры концентрация электронов в зоне проводимости становится сравнимой с концентрацией примеси и выражения (4.111) и (4.112) в этом случае неприменимы. Однако теперь можно рассматривать другой крайний случай, когда температура достаточно высока и выполняется неравенство

$$e^{(E_c - E_d)/kT} \ll N_c/2N_d.$$

При этом (4.110) аппроксимируется выражением

$$F = E_c - kT \ln(N_c/N_d), \quad (4.113)$$

которому соответствует:

$$n_0 = N_d. \quad (4.114)$$

Это означает, что практически вся донорная примесь ионизована и концентрация электронов в зоне проводимости не зависит от тем-

пературы. Эта область температур, при которой имеет место полная ионизация примеси, носит название области истощения примеси (или область полной ионизации примеси) и на рис. 4.14 отмечена цифрой 2.

Условие полной ионизации донорной примеси, когда $n_0 = N_d$, соответствует положению уровня Ферми на несколько kT ниже уровня примеси E_d . Это значит, что при повышении температуры уровень Ферми, понижаясь, пересекает уровень E_d и уходит вниз. Температура, при которой $F = E_d$, носит название температуры истощения T_s , ее можно определить из условия

$$F = E_c - kT_s \ln(N_c/N_d) = E_d, \quad (4.115)$$

откуда

$$T_s = \frac{E_c - E_d}{k \ln(N_c/N_d)}. \quad (4.116)$$

Как следует из (4.116), температура истощения тем ниже, чем меньше энергия ионизации ($E_c - E_d$), и концентрация донорной примеси N_d и чем больше эффективная масса электронов, определяющая величину N_c . При малых значениях ($E_c - E_d$) истощение примеси наступает при очень низких температурах. Например, в электронном германии, легированном сурьмой в количестве $N_d = 10^{16} \text{ см}^{-3}$, для которой энергия ионизации равна 0,0096 эВ, насыщение наступает уже при $T_s = 32 \text{ К}$.

При дальнейшем повышении температуры увеличение концентрации электронов в зоне проводимости будет осуществляться за счет переходов электронов из валентной зоны. В этом случае положение уровня Ферми и концентрация электронов будут определяться уравнениями (4.88), (4.89) и (4.115). На рис. 4.14 область 3 соответствует области собственной проводимости. В этом случае $F = F_i$ и из (4.88) и (4.113) следует, что

$$F_i = E_c - \frac{1}{2} E_g - \frac{kT_i}{2} \ln \frac{N_c}{N_v} = E_c - kT_i \ln \frac{N_d}{N_c}. \quad (4.117)$$

Отсюда получаем

$$T_i = \frac{E_g}{k \ln \left(\frac{N_c N_v}{N_d^2} \right)}. \quad (4.118)$$

Анализ этого выражения показывает, что температура T_i , при которой наступает собственная проводимость у донорного полупроводника, тем ниже, чем меньше ширина запрещенной зоны и концентрация примеси и чем больше значение эффективных масс носителей заряда.

Таким образом, используя описанные приближения, можно проследить изменение концентрации электронов и положения уровня Ферми в запрещенной зоне электронного полупроводника во всей области изменения температуры.

В качестве примера на рис. 4.15, а приведены температурные зависимости уровня Ферми и на рис. 4.15, б — концентрация равно-

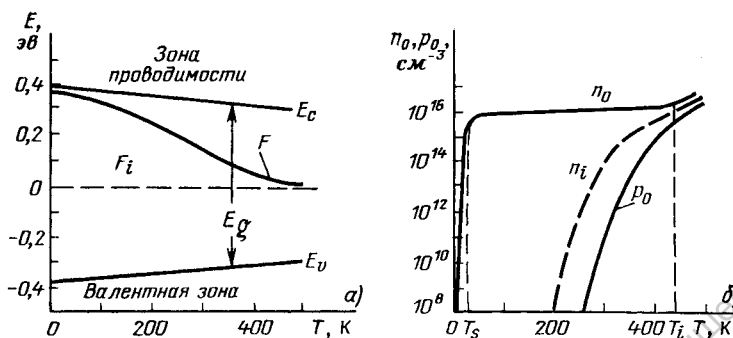


Рис. 4.15. Температурная зависимость уровня Ферми (а) и концентрации носителей заряда (б) для германия, легированного сурьмой

весных электронов n_0 и дырок p_0 для германия, легированного сурьмой в количестве $N_d \approx 10^{16} \text{ см}^{-3}$. Кроме того, на этих кривых пунктиром показан ход F_i и n_i в собственном германии. При построении графиков учтена зависимость ширины запрещенной зоны германия от температуры.

Из этого рисунка следует, что при температуре абсолютного нуля уровень Ферми в германии расположен посередине между дном зоны проводимости E_c и уровнем донорной примеси E_d . При повышении температуры он опускается и приближается к уровню примеси E_d . При температуре насыщения T_s на донорной примеси электроны находятся в количестве, равном:

$$n_d = \frac{N_d}{\frac{1}{2} e^{(E_d - F)/kT} + 1} = \frac{2}{3} N_d,$$

а в зоне проводимости соответственно $1/3 N_d$ электронов. С дальнейшим ростом температуры уровень Ферми продолжает опускаться и наступает область истощения; вся примесь ионизована и концентрация электронов проводимости остается постоянной и равной $n_0 = N_d$. В этой температурной области имеет место уже ионизация атомов основного вещества и появляются неосновные носители заряда — дырки. Их концентрация резко возрастает с ростом температуры согласно соотношению

$$p_0 = \frac{n_i^2}{n_0} = \frac{N_c N_v}{N_d} e^{-E_g/kT}. \quad (4.119)$$

Когда уровень Ферми достигает середины запрещенной зоны, то $n_0 = p_0 = n_i$ и полупроводник от примесного переходит к собственному. При дальнейшем повышении температуры уровень Ферми приближается к той зоне, которая имеет меньшую эффективную плотность состояний.

2. Акцепторный полупроводник ($N_d = 0$)

Определение положения уровня Ферми для невырожденного полупроводника, легированного акцепторной примесью, полностью аналогично случаю донорного полупроводника. Условие электро-нейтральности теперь будет иметь вид:

$$p_0 = n_a = N_a - p_a. \quad (4.120)$$

Отсюда для невырожденного полупроводника с учетом (4.67) и (4.54) легко получить соотношение для

$$p_0 = \frac{N_a}{g e^{-\varepsilon_a - \eta} + 1} = \frac{N_a N_v}{N_v + g e^{\varepsilon_i - \varepsilon_a} p_0}. \quad (4.121)$$

После несложных алгебраических вычислений находим:

$$p_0 = \frac{2N_a}{V \sqrt{1 + 4g e^{\varepsilon_i - \varepsilon_a} N_a N_v^{-1}} + 1}. \quad (4.122)$$

Положение уровня Ферми будет определяться равенством

$$e^{\eta + \varepsilon_i} = \frac{N_v}{p_0} = \frac{N_v}{2N_a} \left[\sqrt{1 + 4g e^{\varepsilon_i - \varepsilon_a} N_a N_v^{-1}} + 1 \right]. \quad (4.123)$$

После логарифмирования получим выражение для уровня Ферми:

$$F = E_c - E_g + kT \ln \left[\frac{N_v}{2N_a} + \sqrt{\left(\frac{N_v}{2N_a} \right)^2 + g N_v N_a^{-1} e^{(E_a - E_v)/kT}} \right]. \quad (4.124)$$

В области низких температур, когда выполняется условие $e^{(E_a - E_v)/kT} \gg \frac{N_v}{2N_a}$, имеем:

$$F = E_c - E_g + \frac{1}{2} (E_a - E_v) + \frac{kT}{2} \ln \left(\frac{g N_v}{N_a} \right) = \frac{E_v + E_a}{2} + \frac{kT}{2} \ln \left(\frac{g N_v}{N_a} \right). \quad (4.125)$$

Таким образом, в дырочном полупроводнике при температуре абсолютного нуля уровень Ферми лежит посередине между потолком валентной зоны и уровнем акцепторной примеси, т. е. $F = \frac{1}{2} (E_v + E_a)$. С увеличением температуры уровень Ферми вначале опускается, когда $N_a > g N_v$, а затем поднимается до уровня E_a . Концентрация свободных дырок в этой области определяется соотношением

$$p_0 = N_v e^{-\eta - \varepsilon_i} = (g^{-1} N_a N_v)^{1/2} e^{(E_v - E_a)/2kT}. \quad (4.126)$$

При высоких температурах, когда $e^{(E_a - E_v)/kT} \ll \frac{N_v}{2N_a}$, из выражения (4.124) получаем:

$$F = E_c - E_g + kT \ln \frac{N_v}{N_a}, \quad (4.127)$$

а концентрация свободных дырок

$$p_0 = N_v e^{-\ln(N_v/N_a)} = N_a. \quad (4.128)$$

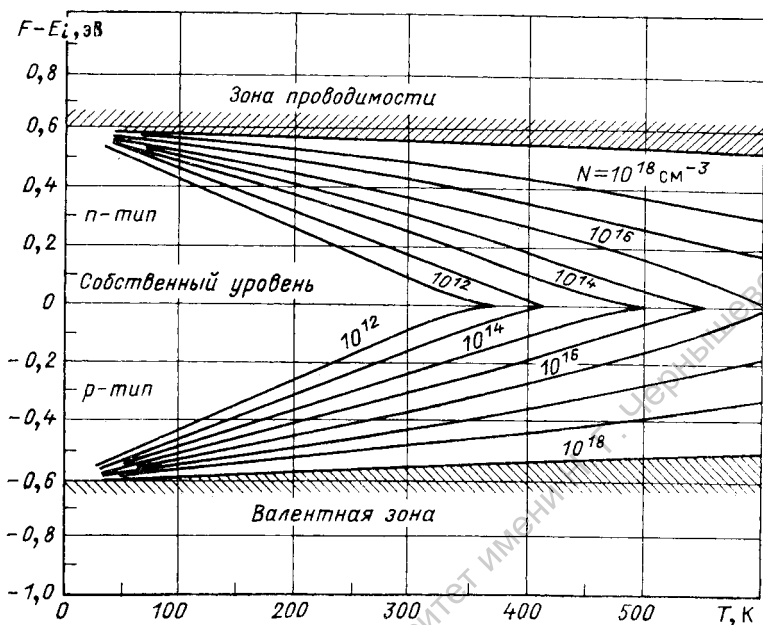


Рис. 4.16. Зависимость уровня Ферми в кремнии от температуры и концентрации примесей

В акцепторном полупроводнике, как и в случае донорной примеси, при высоких температурах наступает область истощения, характеризующаяся полной ионизацией атомов акцепторной примеси. С дальнейшим ростом температуры уровень Ферми поднимается к середине запрещенной зоны и полупроводник ведет себя как собственный.

Уровень Ферми для кремния в зависимости от концентрации примесей и температуры показан на рис. 4.16. Здесь же приведена и зависимость ширины запрещенной зоны от температуры.

4.8. ЗАВИСИМОСТЬ ПОЛОЖЕНИЯ УРОВНЯ ФЕРМИ ОТ ТЕМПЕРАТУРЫ ДЛЯ НЕВЫРОЖДЕННОГО ПОЛУПРОВОДНИКА С ЧАСТИЧНО КОМПЕНСИРОВАННОЙ ПРИМЕСЬЮ

Рассмотренные случаи чисто донорного и чисто акцепторного полупроводника обычно не реализуются на практике. Как бы тщательно ни проводилась очистка материала, в нем обычно остаются примеси нескольких видов, причем некоторые из примесных центров могут отдавать или присоединять к себе более одного электрона.

Остановимся более подробно на случае, когда наряду с одно-электронной донорной примесью в полупроводнике имеется некоторое количество акцепторной примеси (см. рис. 4.13).

Предположим, что в полупроводнике имеется в равных количествах донорная и акцепторная примеси, т. е. $N_d = N_a$. При температуре абсолютного нуля электроны занимают наиболее низкие энергетические состояния. Это означает, что в зоне проводимости нет электронов, а в валентной зоне все состояния заняты. Поскольку есть свободные состояния на акцепторной примеси, которые энергетически расположены ниже, чем состояния донорной примеси, то электроны, стремясь заполнить состояния с наименьшей энергией, перейдут с донорных уровней на акцепторные. В результате образуются в равных количествах положительные и отрицательные ионы. Условие электронейтральности выполняется, поскольку $N_d^+ = N_a^-$. Такой полупроводник называют **компенсированным**. В нем имеет место взаимная компенсация примесей, вследствие чего последние не могут являться поставщиками электронов и дырок, носители заряда возникают в результате тепловой генерации «зона — зона».

В компенсированном полупроводнике из-за наличия примесей периодичность поля решетки нарушена значительно сильнее, чем в истинно собственном. Это будет сказываться на эффектах, связанных с рассеянием носителей заряда, в частности на их подвижности.

Будем теперь считать, что $N_d > N_a$ и во всей рассматриваемой области температур уровень Ферми лежит значительно выше E_a , т. е. лежит выше уровня энергии акцепторной примеси. При этом условии все акцепторные состояния заполнены электронами, перешедшими с уровней донорной примеси. В результате этого общее количество электронов на донорных уровнях и в зоне проводимости составляет $N_d - N_a$. Пока не наступит собственная проводимость, свободные дырки в валентной зоне отсутствуют и полупроводник ведет себя как чисто донорный. Такой же полупроводник называют **частично компенсированным**. Степень компенсации определяется отношением N_a/N_d . В таком полупроводнике перенос заряда осуществляется только электронами, и постоянная Холла имеет отрицательный знак.

В случае, когда $N_a > N_d$ и уровень Ферми проходит значительно ниже E_d , донорная примесь полностью ионизована. Все электроны с донорных уровней перешли на уровни акцепторной примеси, поэтому при $T = 0$ на атомах акцепторной примеси будет $N_a - N_d$ дырок. С ростом температуры электроны из валентной зоны будут переходить к атомам акцепторной примеси, полупроводник будет вести себя как дырочный, имеющий положительный коэффициент Холла. С наступлением собственной проводимости знак постоянной Холла изменится, так как подвижность электронов больше подвижности дырок.

Определим зависимость концентрации свободных электронов и положения уровня Ферми от температуры в полупроводнике с частично компенсированной примесью для случая $N_d > N_a$. Условие электронейтральности при ионизации только примеси имеет вид:

$$n_0 + n_d = N_d - N_a \quad (4.129)$$

или

$$n_0 + N_a = p_d. \quad (4.130)$$

Будем считать, что полупроводник невырожден. Используя соотношение $n_0 = N_c e^n$, получаем из (4.52) и (4.130):

$$n_0 + N_a = \frac{N_d N_c}{N_c + g e^{\varepsilon_d} n_0}. \quad (4.131)$$

Решив это уравнение, найдем:

$$n_0 = \frac{2(N_d - N_a)}{1 + \frac{g N_a}{N_c} e^{\varepsilon_d} + \left[\left(1 + \frac{g N_a}{N_c} e^{\varepsilon_d} \right)^2 + \frac{4g(N_d - N_a) e^{\varepsilon_d}}{N_c} \right]^{1/2}}. \quad (4.132)$$

Из соотношения $e^n = n_0/N_c$ определим энергию Ферми

$$F = E_c + kT \times \\ \times \ln \frac{2(N_d - N_a) N_c^{-1}}{1 + \frac{g N_a}{N_c} e^{(E_c - E_d)/kT} + \left[\left(1 + \frac{g N_a}{N_c} e^{(E_c - E_d)/kT} \right)^2 + \right. \\ \left. + \frac{4g(N_d - N_a) e^{(E_c - E_d)/kT}}{N_c} \right]^{1/2}}}. \quad (4.133)$$

Нетрудно видеть, что при $N_a = 0$ найденные выражения совпадают с формулами (4.107) и (4.110) для невырожденного донорного полупроводника.

В области высоких температур, когда донорные атомы полностью ионизованы, справедливы неравенства:

$$\frac{g N_a}{N_c} e^{\varepsilon_d} \ll 1 \quad \text{и} \quad \frac{4g(N_d - N_a) e^{\varepsilon_d}}{N_c} \ll 1. \quad (4.134)$$

Учет этих неравенств дает следующие соотношения:

$$n_0 = N_d - N_a; \quad (4.135)$$

$$F = E_c + kT \ln \frac{N_d - N_a}{N_c}. \quad (4.136)$$

Как и следовало ожидать, концентрация свободных электронов постоянна и определяется количеством некомпенсированной донорной примеси. Уровень Ферми с ростом температуры понижается и при температурах, соответствующих началу собственной проводимости, приближается к середине запрещенной зоны.

В области низких температур вид выражений для n_0 и F зависит от соотношения концентраций доноров и акцепторов. Если степень компенсации мала, т. е. $N_a \ll N_d$, то можно найти область температур, для которой выполняется соотношение $N_a \ll n_0 \ll N_d$, что эквивалентно условию

$$\frac{g N_a}{N_c} e^{\varepsilon_d} \ll 1 \ll \frac{4g(N_d - N_a) e^{\varepsilon_d}}{N_c}. \quad (4.137)$$

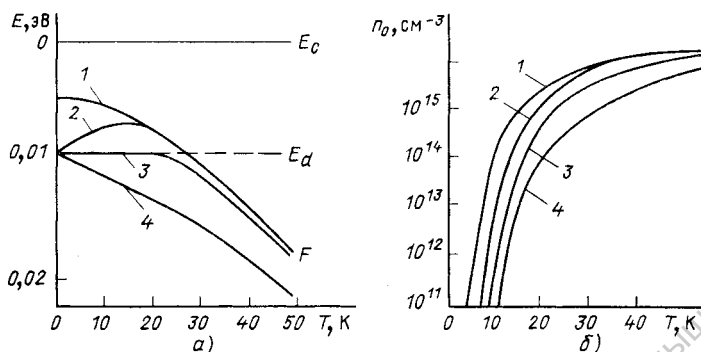


Рис. 4.17. Зависимость положения уровня Ферми (а) и концентрации электронов (б) от температуры в германии при различной степени компенсации. Для всех кривых $N_d - N_a = 10^{16} \text{ см}^{-3}$. N_a равно:

1 — 0; 2 — 10^{14} ; 3 — 10^{15} ; 4 — 10^{16} см^{-3}

В этой области температурные зависимости концентрации и уровня Ферми на основании (4.132) и (4.133) приближенно описываются формулами:

$$n_0 \approx (g^{-1} N_c N_d)^{1/2} e^{-(E_c - E_d)/2kT}; \quad (4.138)$$

$$F = \frac{E_c + E_d}{2} + \frac{kT}{2} \ln \frac{N_d}{gN_c}, \quad (4.139)$$

совпадающими с выражениями (4.112) и (4.111) для донорного полупроводника. Понятно, что при $N_d \gg N_a$ концентрация свободных электронов n_0 не зависит от количества акцепторных атомов, как и следует из равенства (4.138). При $T = 0$ уровень Ферми лежит в середине между E_c и E_d . С ростом температуры он понижается (кривая 1 на рис. 4.17, а).

Однако при более низких температурах, когда $n_0 \ll N_a$, влиянием акцепторов уже нельзя пренебречь. Положение уровня Ферми в этой области существенно отличается от того, что имеет место в чисто донорном полупроводнике [соотношение (4.111)].

В случае сильной компенсации справедливо неравенство $n_0 \ll N_a < N_d$, поэтому $\frac{gN_a}{N_c} e^{E_d} \gg 1$ и $\frac{gN_a}{N_c} e^{E_d} > \frac{4g(N_d - N_a)e^{E_d}}{N_c}$ и из выражений (4.132) и (4.133) следует:

$$n_0 \approx N_c \frac{N_d - N_a}{gN_a} e^{-(E_c - E_d)/kT}; \quad (4.140)$$

$$F = E_d + kT \ln \frac{N_d - N_a}{gN_a} = E_d + kT \ln \frac{N_d - N_a}{2N_a}, \quad (4.141)$$

так как $g = 2$.

В этом случае энергия активации равна энергии ионизации донорной примеси $w = E_c - E_d$, т. е. в 2 раза больше, чем у чисто донорного полупроводника.

На основании соотношения (4.141) на рис. 4.17, а построена зависимость $F(T)$, а на рис. 4.17, б изображена зависимость $n_0(T)$ для различной степени компенсации. Видно, что при повышении температуры, начиная от абсолютного нуля, если $\frac{N_d - N_a}{2N_a} > 1$, уровень Ферми вначале повышается от уровня донорной примеси, а затем понижается (кривая 2). В противоположном случае, когда $\frac{N_d - N_a}{2N_a} < 1$, с ростом температуры уровень Ферми снижается от донорного уровня (кривая 4), а при $\frac{N_d - N_a}{2N_a} = 1$ он остается неизменным, пока выполняется условие $n_0 \ll N_a$, а затем понижается (кривая 3). Из сравнения кривых, представленных на рис. 4.17, б, следует, что температура, при которой наступает насыщение зависимости $n_0(T)$, растет с увеличением степени компенсации донорной примеси.

4.9. ПРИМЕСНЫЕ ПОЛУПРОВОДНИКИ ПРИ ОЧЕНЬ НИЗКИХ ТЕМПЕРАТУРАХ

Отметим еще одну особенность примесных полупроводников. У донорного полупроводника при температурах, недостаточных для ионизации основного вещества, наличие электронов в зоне проводимости определяется процессом теплового возбуждения атомов донорной примеси. При этом количество электронов проводимости равно количеству образовавшихся положительных ионов примеси, т. е.

$$n_0 = p_d = \frac{N_d}{g e^{(F - E_d)/kT} + 1}. \quad (4.142)$$

В таком полупроводнике при температуре, когда kT меньше энергии ионизации донора или акцептора, удельное сопротивление $\rho = (en_0\mu)^{-1}$ и коэффициент Холла ($R = -1/en_0$) примесного полупроводника согласно (4.142) должны возрастать экспоненциально при понижении температуры. Однако, как показывают исследования, у полупроводников с концентрацией примеси свыше 10^{17} см^{-3} коэффициент Холла при низких температурах проходит через максимум. Положение этого максимума, как видно из рис. 4.18, зависит от концентрации примеси. Температурная зависимость удельного сопротивления этих образцов приведена на рис. 4.19. У них при температуре, когда имеет место максимум коэффициента Холла, удельное сопротивление достигает насыщения, возрастая при дальнейшем понижении температуры значительно медленнее. У образцов с содержанием примеси свыше 10^{19} см^{-3} удельное сопротивление не зависит от температуры.

Из этих экспериментальных фактов следует, что в полупроводнике с большим содержанием примеси при низких температурах имеет место явление, получившее название **проводимости примесной зоны**. При этом механизм примесной про-

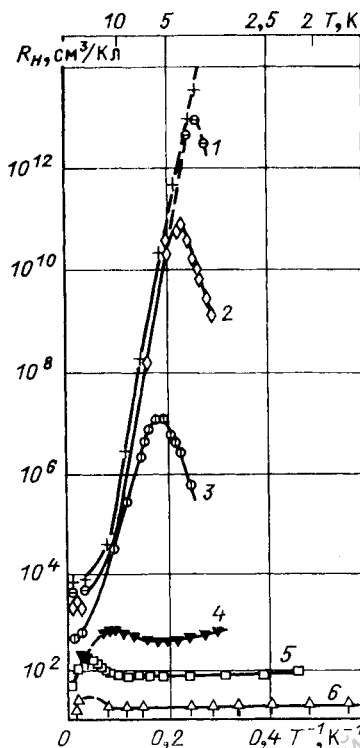


Рис. 4.18. Коэффициент Холла дырочных образцов германия как функция температуры при низких ее значениях. Концентрация бора: 1 — $6,7 \cdot 10^{17}$; 2 — $1,3 \cdot 10^{18}$; 3 — $2,7 \cdot 10^{18}$; 4 — $6,7 \cdot 10^{18}$; 5 — $1,3 \cdot 10^{19}$; 6 — $2,7 \cdot 10^{19}$ см⁻³

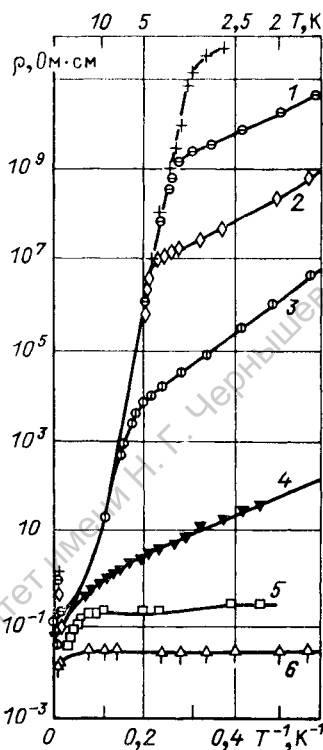


Рис. 4.19. Удельное сопротивление дырочных образцов германия как функция температуры при низких ее значениях. Концентрации примеси те же, что и на рис. 4.18

водимости различен в зависимости от концентрации введенной примеси.

Водородоподобная модель примесных атомов не учитывает взаимодействия их между собой. Однако при достаточно высоких концентрациях атомов примеси учет их взаимодействия становится необходимым, поскольку боровский радиус для электрона, связанного на водородоподобном примесном центре, согласно (2.175) пропорционален относительной диэлектрической проницаемости и обратно пропорционален эффективной массе электрона. В силу этого эффективный радиус орбиты для электрона примеси в германии примерно в 60 раз больше, чем радиус Бора у электрона в атоме водорода. Поэтому уже при концентрации примеси примерно 10^{16} см⁻³ волновые функции электронов, связанных на нейтральной донорной примеси, значительно перекрываются. В этом случае примесный уровень, расплываясь, превращается в примесную зону, располо-

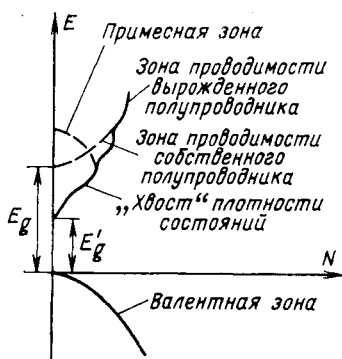


Рис. 4.20. Плотность состояний для вырожденного полупроводника

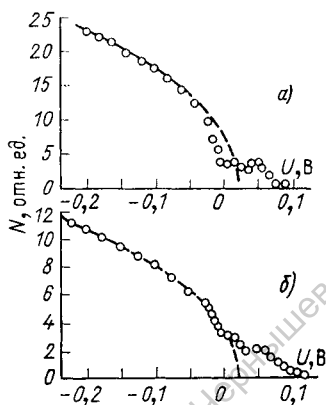


Рис. 4.21. Плотность состояний N в арсениде галлия p -типа:
 а — $N_d = 5,4 \cdot 10^{18} \text{ см}^{-3}$; б — $N_d = 9,9 \cdot 10^{18} \text{ см}^{-3}$ (пунктиром показана наилучшая аппроксимация к параболической зоне)

женную вблизи дна зоны проводимости. Ширина этой энергетической полосы зависит от степени перекрытия волновых функций электронов атомов примеси, а следовательно, зависит от концентрации донорной примеси. В этом случае возникает проводимость по примесной зоне, обусловленная электронами донорной примеси, которые ведут себя подобно электронному газу.

При больших концентрациях донорной примеси примесная зона, расширяясь, может перекрываться с зоной проводимости (или с валентной зоной в случае акцепторного полупроводника). Таким образом, в кристалле образуется единая зона разрешенных значений энергии.

Наличие примесной зоны влияет на плотность состояний в зоне проводимости вблизи уровня E_c (или соответственно вблизи уровня E_v). При концентрации примеси порядка 10^{19} см^{-3} перекрытие зон невелико и зависимость плотности состояний от энергии вблизи уровня E_c (или E_v) сильно искажается, а зона проводимости (или валентная зона) становится нестандартной, как это показано на рис. 4.20. У такого полупроводника вследствие высокой степени легирования образуется так называемый «хвост» плотности состояний и ширина запрещенной зоны E_g уменьшается до E'_g . У вырожденного германия со степенью легирования 10^{19} см^{-3} сужение ширины запрещенной зоны составляет несколько сотых долей электрон-вольта. С увеличением степени легирования эффект «хвоста» плотности состояний возрастает. Это следует из рис. 4.21, на котором даны результаты исследования по определению плотности состояний в арсениде галлия p -типа, полученные при изучении туннелирования в контакте металл — полупроводник.

Несмотря на то что вырожденные полупроводники уже нашли весьма широкое применение в некоторых полупроводниковых приборах (например, туннельные диоды, квантовые генераторы), строгая теория сильно легированных полупроводников является математически сложной задачей.

Эксперименты показали, что у полупроводников n - или p -типа при низких температурах электропроводность может наблюдаться также при концентрации примеси 10^{16} см⁻³. Этот механизм электропроводности получил название прыжкового. Рассмотрим его на примере электронного полупроводника. В реальном донорном полупроводнике всегда имеется некоторое количество акцепторных центров N_a , поэтому часть электронов с уровней донорной примеси N_d перейдет на уровни акцепторов. В результате такой компенсации даже при очень низкой температуре доноры в количестве N_d оказываются ионизованными. Поэтому электрон, локализованный на одном из доноров, может «перескочить» на ближайший ионизованный донор, ионизировав при этом первый и нейтрализовав последний. Такой перескок электрона происходит случайным образом. При наличии внешнего электрического поля прыжки электронов приобретут направленный характер, в результате чего возникнет поток заряда. Так как доноры расположены на больших расстояниях друг от друга, то подвижность носителей заряда, перемещающихся за счет перескокового механизма, значительно меньше подвижности электронов в зоне проводимости. В области же высоких температур будет преобладать обычная проводимость по зонам энергии. Поэтому для всей области изменения температуры проводимость σ и коэффициент Холла R такого образца будут определяться выражениями:

$$\sigma = en_1\mu_1 + en_2\mu_2; \quad (4.143)$$

$$R = \frac{r_1 n_1 \mu_1^2 + r_2 n_2 \mu_2^2}{e(n_1 \mu_1 + n_2 \mu_2)^2} = \frac{n_1 \mu_1 \mu_{1H} + n_2 \mu_2 \mu_{2H}}{e(n_1 \mu_1 + n_2 \mu_2)^2}, \quad (4.144)$$

где согласно соотношению (6.115) $r_1 = \frac{\mu_{1H}}{\mu_1}$; $r_2 = \frac{\mu_{2H}}{\mu_2}$, а n_1, μ_1, μ_{1H} и n_2, μ_2, μ_{2H} — концентрации, дрейфовые и холловские подвижности электронов в зоне проводимости и в области примесной проводимости.

4.10. НЕКРИСТАЛЛИЧЕСКИЕ ПОЛУПРОВОДНИКИ

В аморфных полупроводниках, которые не являются кристаллическими, сохраняется только ближний порядок, так что в пределах нескольких координационных сфер расположение атомов в кристалле и аморфном материале практически идентично. Но по мере удаления от атома, принятого за начало отсчета, влияние случайного фактора в аморфном материале возрастает и между двумя атомами, находящимися на большом расстоянии друг от друга, не удастся установить строгих геометрических соотношений, характерных для кристалла. Поскольку в некристаллических материалах окружение атома случайным образом меняется от атома к атому, а следовательно, для

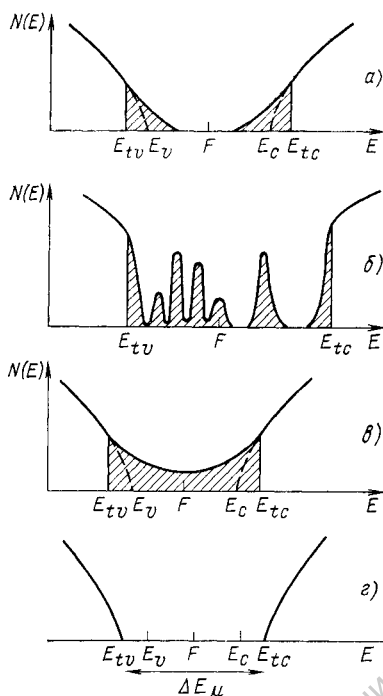


Рис. 4.22. Схема изменения зон в аморфном полупроводнике:

а — хвосты зон, состоящие из локализованных состояний (заштрихованные области); *б* — возможное распределение плотности локализованных состояний; *в* — перекрытие хвостов локализованных состояний, приводящее к исчезновению запрещенной зоны; *г* — запрещенная зона подвижности

локализованными состояниями, заштрихованы. Кроме того, плотность состояний может проходить через многочисленные максимумы, заметно отдаленные друг от друга (рис. 4.22, б). Для распределения $N(E)$, представленного на рис. 4.22, в, «хвосты» зон перекрываются таким образом, что во всем промежутке между E_{tc} и E_{tv} существуют уровни энергии локализованных состояний. Поскольку электропроводность по локализованным состояниям запрещена, то подвижность носителей заряда, находящихся в этих состояниях, равна нулю. Поэтому для аморфных материалов термин «запрещенная зона» не соответствует энергетической структуре уровней и принято говорить о запрещенной зоне подвижности («зазоре» подвижности), значение которой $\Delta E_{\mu} = E_{tc} - E_{tv}$ (рис. 4.22, г) превышает ширину запрещенной зоны $E_g = E_c - E_v$ в кристаллическом материале.

них характерно отсутствие дальнего порядка в расположении атомов, поэтому для аморфного материала нельзя использовать представление о средней длине свободного пробега, а также не применимо представление о зонах Бриллюэна. Однако, как показано теоретически, в некристаллических материалах могут существовать состояния электронов, волновые функции которых простираются на область макроскопических размеров. Эти состояния, получившие название «расширенных» состояний, могут образовывать энергетические зоны, похожие на энергетические зоны в кристалле. Кроме того, в аморфных материалах существуют связанные состояния, волновая функция которых локализована в области с размерами ненамного больше межатомного расстояния. Уровни энергии, соответствующие локализованным состояниям, расположены вблизи краев зон. На рис. 4.22, а, на котором представлена плотность состояний $N(E)$, энергии E_c и E_v , соответствуют краям зон в кристаллическом материале, а энергии E_{tc} и E_{tv} разделению между разрешенными и связанными состояниями. Области, занятые ло-

Экспериментальные работы в области аморфных полупроводников выявили тесную связь электрических и оптических свойств этих материалов с распределением локализованных состояний, а также была показана тесная связь между плотностью состояний в запрещенной зоне подвижности и технологией изготовления материала. Примером влияния технологии приготовления материала могут служить данные рис. 4.23 для пленок аморфного кремния (а — Si), полученных методом тлеющего разряда (кривая 1) и методом вакуумного напыления (кривая 2). Стрелкой указано положение уровня Ферми для нелегированного а — Si. Этот энергетический спектр

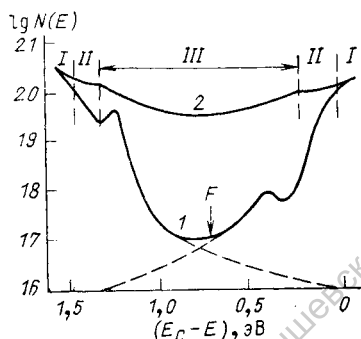


Рис. 4.23. Распределение плотности состояний в аморфных пленках кремния, полученных методом тлеющего разряда (1) и методом напыления в вакууме (2)

по характеру электронных состояний может быть разделен на три области: область I — нелокализованные состояния; область II — локализованные состояния в хвосте зон и область III — локализованные состояния в зазоре подвижности. Из графиков следует, что при изменении технологии изготовления пленок а — Si плотность локализованных состояний в зазоре подвижности может изменяться на несколько порядков. Эти состояния нейтральны, когда они не заполнены электронами. Ниже уровня Ферми F эти состояния захватывают электроны и заряжены отрицательно. Выше уровня Ферми эти состояния заряжены положительно.

Для большинства аморфных материалов выполняется правило, согласно которому в аморфном материале сохраняется первое координационное число соответствующего кристалла. Это фундаментальное правило может быть использовано для классификации аморфных материалов.

В табл. 4.1 приведено распределение ряда аморфных материалов в соответствии с их координационным числом.

Таблица 4.1

Число ближайших соседей	Группа в периодической системе	Материал
2	A^{VI}	Se, Te
3	A^V	As, Sb, P
3 и 4	A^{IV}	C
4	A^{IV}	Ge, Si
6	A^{III}	B
3—2	$A^V B^{VI}$	$(As, Sb, Bi)_2(S, Se, Te)_3$
4—2	$A^{IV}-B^{VI}$	$SiO_2, (Ge, Sn, Pb)(S, Se, Te)$

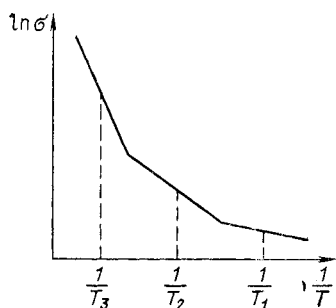


Рис. 4.24. Температурная зависимость удельной проводимости в аморфных материалах

Как следует из данных таблицы, аморфные материалы даже в первой координационной сфере могут обладать большим структурным разнообразием. И все-таки по структурному признаку их можно разделить на два класса. Это — материалы с тетраэдрической координацией и материалы типа халькогенидных стекол. Для материалов первой группы характерен жесткий структурный каркас с весьма ограниченной возможностью для локальной атомной перестройки. Материалы второй группы, как правило, характеризуются меньшим

числом ближайших соседей и обладают большей податливостью. Для тетраэдрических материалов типа а — Si характерно большое количество оборванных связей (обычно их называют «болтающимися» связями). Наличие электронов с неспаренными спинами приводит к появлению сигнала ЭПР, который для плотности парамагнитных центров дает концентрацию $10^{19} - 10^{20} \text{ см}^{-3}$. Введение в эти материалы водорода приводит к насыщению водородных связей и исчезновению сигнала ЭПР. Такой материал удастся легировать как донорной, так и акцепторной примесью, и он используется для получения некоторых полупроводниковых приборов.

Материалы второй группы (типа селена и халькогенидных стекол) не дают сигнала ЭПР, что свидетельствует об отсутствии парамагнитных центров или по крайней мере о том, что их плотность ниже обнаруживаемого предела 10^{15} см^{-3} . Имеющиеся в этом случае оборванные связи должны быть заполнены двумя электронами с антипараллельными спинами.

Другая особенность халькогенидных материалов состоит в том, что они имеют на внешней оболочке шесть электронов и не все участвуют в образовании связи. Так, например, для Se в образовании связи участвуют два электрона, четыре электрона находятся в несвязанном состоянии, так называемые неподеленные пары. Именно электроны неподеленных пар и образуют валентную зону халькогенидных стеклообразных полупроводников. Наличие неподеленных пар приводит к особенностям энергетической структуры и физических свойств стеклообразных материалов.

В аморфных полупроводниках при температуре, отличной от нуля, электропроводность может осуществляться прыжковым механизмом по локализованным состояниям. Поскольку вероятность перескока электронов от одного атома к другому зависит от температуры, эта проводимость будет термоактивированной (рис. 4.24). Характер температурной зависимости будет определяться тем, по какой группе уровней при данной температуре происходит перескок электронов и каков его механизм.

Оптические свойства аморфных образцов в значительно большей степени зависят от условий их приготовления, чем свойства кристаллов. Это объясняется не только тем, что с изменением условий приготовления аморфного материала может значительно модифицироваться кривая его плотности состояний, но также и тем, что могут возникать микропоры (особенно в тетраэдрических материалах), приводящие к рассеянию света в образце.

В стеклообразных полупроводниках наблюдается эффект токовой неустойчивости, на основе которого можно создавать переключатели, запоминающие устройства, имеющие очень малые времена переключения — порядка 10^{-10} с.

ГЛАВА ПЯТАЯ

РАССЕЯНИЕ ЭЛЕКТРОНОВ И ДЫРОК В ПОЛУПРОВОДНИКАХ

5.1. МЕХАНИЗМЫ РАССЕЯНИЯ ЭЛЕКТРОНОВ И ДЫРОК

При рассмотрении движения носителей заряда в идеальном периодическом поле кристаллической решетки было установлено, что электроны и дырки движутся как свободные частицы (см. § 2.10). В реальном кристалле электроны и дырки совершают сложные траектории движения вследствие соударений с дефектами решетки. Поскольку дефекты, искажающие периодичность поля решетки и являющиеся центрами рассеяния, имеют разную природу, то они будут обуславливать и различные механизмы рассеяния носителей заряда. *В полупроводниках центрами рассеяния могут быть тепловые колебания решетки и статические дефекты, такие как атомы и ионы примеси, вакансии, дислокации, границы двойников и кристаллитов.* Для количественной оценки процесса рассеяния вводится параметр σ , называемый эффективным сечением рассеяния.

Предположим, что имеется n свободных электронов, которые со средней тепловой скоростью v_0 движутся в данном направлении. Тогда nv_0 есть плотность потока электронов, т. е. количество электронов, проходящих в единицу времени через единичную площадку образца, перпендикулярную направлению их скорости. Допустим, что на пути потока электронов в единичном сечении образца имеется N одинаковых центров рассеяния. Каждый центр характеризуется эффективным сечением, равным σ . Это, по существу, то пространство вокруг центра, в области которого имеет место рассеяние электронов. Поэтому количество рассеянных электронов в единицу времени n_1 определяется эффективным сечением σ , количеством центров рассеяния N и плотностью падающего потока электронов nv_0 , т. е.

$$n_1 = \sigma N n v_0. \quad (5.1)$$

Если W — вероятность рассеяния одной частицы в единицу времени, то количество рассеянных электронов в единицу времени n_1 есть

$$n_1 = W n. \quad (5.2)$$

Тогда на основании (5.1) и (5.2) можно написать

$$\sigma = \frac{n_1/N}{nv_0} = \frac{W}{Nv_0}. \quad (5.3)$$

Таким образом, *эффективное сечение рассеяния σ есть отношение числа электронов, удаленных из пучка в результате рассеяния на одном центре в единицу времени, к плотности падающего пучка частиц*. Эффективное сечение рассеяния имеет размерность площади

$$[\sigma] = \frac{[W]}{[Nv_0]} = \frac{T^{-1}}{L^{-3}LT^{-1}} = L^2.$$

Из формулы (5.3) найдем, что вероятность рассеяния

$$W = \sigma Nv_0. \quad (5.4)$$

Следовательно, *вероятность рассеяния или вероятность столкновения определяется эффективным сечением, количеством центров рассеяния и скоростью движения носителя заряда*.

В то же время вероятность столкновения обратно пропорциональна времени свободного пробега:

$$W = 1/\tau. \quad (5.5)$$

Поэтому время свободного пробега и длину свободного пробега можно выразить через эффективное сечение. Из (5.4) и (5.5) следует, что

$$\tau = \frac{1}{\sigma Nv_0} = \frac{l}{v_0} \quad (5.6)$$

или

$$l = \frac{1}{\sigma N}. \quad (5.7)$$

Величина $l^{-1} = \sigma N$ есть *вероятность рассеяния на единичном интервале пути*.

Рассмотрим случай, когда имеются различные центры рассеяния. Пусть i -й центр характеризуется эффективным сечением σ_i и число таких центров равно N_i . Механизм рассеяния на этих центрах определяет длину свободного пробега l_i . Если все процессы рассеяния возможны, то согласно теории вероятности полная вероятность рассеяния в единицу времени будет определяться суммой отдельных вероятностей рассеяния

$$W = \sum_i W_i. \quad (5.8)$$

Так как $\sum_i \sigma_i N_i = \sum_i l_i^{-1}$, то полная длина свободного пробега согласно (5.4) и (5.8) может быть определена из соотношения

$$l^{-1} = \sum_i l_i^{-1}. \quad (5.9)$$

Согласно (5.9) *полная длина свободного пробега всегда меньше самой малой парциальной длины свободного пробега*.

Поскольку роль дефектов в процессе рассеяния носителей заряда различна, поэтому различные дефекты должны иметь разное эффективное сечение. Для их количественной оценки за эффективное сечение рассеяния σ примем площадку, в пределах которой возможно взаимодействие между носителем заряда и дефектом.

Такие дефекты, как вакансии, междоузельные атомы, во многих отношениях сходны с примесями замещения. Эти дефекты называются точечными дефектами. Для них за σ можно принять площадь квадрата со стороной, равной постоянной решетки, т. е. $\sigma_A = (5 \cdot 10^{-8})^2 \approx 3 \cdot 10^{-15} \text{ см}^2$. Если предположить, что концентрация атомов примеси $N_A = 10^{16} \text{ см}^{-3}$, то длина свободного пробега при рассеянии носителей заряда на атомах примеси будет составлять $l_A = (3 \cdot 10^{-15} \cdot 10^{16})^{-1} \approx 3 \cdot 10^{-2} \text{ см} = 300 \text{ мкм}$.

Для иона примеси будем считать, что его диаметр в 10 раз больше диаметра примесного центра, т. е. $\sigma_I = (5 \cdot 10^{-8} \cdot 10)^2 \approx 3 \cdot 10^{-13} \text{ см}^2$. В случае, если $N_I = 10^{16} \text{ см}^{-3}$, то $l_I = 3 \cdot 10^{-4} \text{ см} = 3 \text{ мкм}$.

Дислокации являются линейными дефектами, простирающимися на большие области кристалла. Предположим, что линейная дислокация имеет длину 0,1 см, а ее диаметр измеряется сотней периодов решетки. В этом случае площадь ее осевого сечения равна $5 \cdot 10^{-8} \times 100 \cdot 10^{-1} = 5 \cdot 10^{-7} \text{ см}^2$. При объемной плотности дислокаций $N_D = 10^8 \text{ см}^{-3}$ длина свободного пробега будет порядка $l_D = 2 \cdot 10^{-2} \text{ см} = 200 \text{ мкм}$.

Границы двойников и кристаллитов, а также дефекты упаковки представляют собой двумерные нерегулярности. Они имеют место в блочных кристаллах и поликристаллических образцах.

Эффективное сечение рассеяния на тепловых колебаниях решетки определяется площадью сечения области, которую занимает колеблющийся атом за вычетом площади сечения самого атома. Если считать, что амплитуда колебаний порядка $r = 0,05 \text{ нм} = 5 \cdot 10^{-9} \text{ см}$, а диаметр атома $d = 10^{-8} \text{ см}$, то $\sigma_T = (d + r)^2 - d^2 \approx 2rd = 10^{-16} \text{ см}^2$. Это значение много меньше, чем для других центров рассеяния, но число колеблющихся атомов велико: $N_T = 10^{22} \text{ см}^{-3}$, поэтому $l_T = 10^{-6} \text{ см} = 0,01 \text{ мкм}$.

5.2. КИНЕТИЧЕСКОЕ УРАВНЕНИЕ БОЛЬЦМАНА

До сих пор совокупность носителей заряда в кристалле рассматривалась как система частиц, находящихся в термодинамическом равновесии. В этом случае движение носителей заряда хаотично, их поведение описывается равновесной функцией распределения, зависящей в общем случае от энергии частиц:

$$f_0 = \frac{1}{e^{(E-F)/kT} + 1}. \quad (5.10)$$

Если поместить кристалл во внешнее постоянное электрическое поле, то движение носителей заряда приобретает упорядоченный характер.

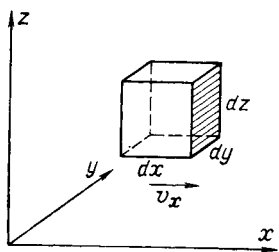


Рис. 5.1. Элементарный объем в геометрическом пространстве

Явления, обусловленные движением носителей заряда под действием электрического и магнитного полей, градиента температуры и градиента концентрации носителей заряда, называют явлениями переноса или кинетическими явлениями.

К явлениям переноса относятся электропроводность, электронная теплопроводность, гальваномагнитные, термоэлектрические и термомагнитные эффекты.

При наличии внешних сил носители заряда находятся в неравновесных условиях. Естественно, что в неравновесном состоянии система носителей заряда описывается неравновесной функцией распределения, которая согласно (2.102) и рис. 2.13 должна зависеть как от энергии частиц, т. е. от $\mathbf{k}$, так и от их координаты $\mathbf{r}$ и времени t , следовательно, функцией $f(\mathbf{k}, \mathbf{r}, t)$.

Для того чтобы определить концентрацию электронов в неравновесном состоянии, поступим следующим образом. В фазовом пространстве кристалла единичного объема, т. е. в пространстве координат x, y, z и квазиимпульсов p_x, p_y, p_z , выделим элемент объема $d\Gamma$. Это будет пространство, построенное вокруг точки с координатами (x, y, z, p_x, p_y, p_z) :

$$d\Gamma = dx dy dz dp_x dp_y dp_z = d\tau_r d\tau_p = \hbar^3 d\tau_k d\tau_r, \quad (5.11)$$

где

$$d\tau_r = dx dy dz; \quad d\tau_p = dp_x dp_y dp_z = \hbar^3 dk_x dk_y dk_z = \hbar^3 d\tau_k.$$

В элементе объема фазового пространства кристалла единичного объема имеется $d\Gamma/\hbar^3$ фазовых ячеек, в каждой из которых могут быть два электрона с антипараллельными спинами. Следовательно, в элементе объема $d\Gamma$ фазового пространства содержится $2d\Gamma/\hbar^3$ квантовых состояний. Если $f(\mathbf{k}, \mathbf{r}, t)$ есть вероятность нахождения электрона в этих состояниях, то число электронов в элементе фазового пространства в момент времени t будет равно:

$$dn = 2 \frac{d\Gamma}{\hbar^3} f(\mathbf{k}, \mathbf{r}, t) = \frac{1}{4\pi^3} f(\mathbf{k}, \mathbf{r}, t) d\tau_k d\tau_r. \quad (5.12)$$

Проинтегрировав выражение (5.12) по всему фазовому пространству V_Γ , получим полное число электронов:

$$n = \frac{1}{4\pi^3} \int_{(V_\Gamma)} f(\mathbf{k}, \mathbf{r}, t) d\tau_k d\tau_r. \quad (5.13)$$

Таким образом, при рассмотрении кинетических явлений необходимо знание неравновесной функции распределения. Выведем уравнение, которому удовлетворяет неравновесная функция распределения $f(\mathbf{k}, \mathbf{r}, t)$. Для этого рассмотрим движение электронов в обычном пространстве и пространстве волновых факторов.

Сначала рассмотрим систему электронов в обычном, геометрическом пространстве. Предположим, что внешнее электрическое поле приложено таким образом, что электроны движутся со скоростью v_x вдоль положительного направления оси x . Определим изменение концентрации электронов за время dt внутри элемента объема $d\tau_r$.

В элемент объема $d\tau_r$, который изображен на рис. 5.1, через его левую грань за время dt согласно (5.2) войдет $dN_{\text{л}}$ электронов, движущихся со скоростью v_x :

$$dN_{\text{л}} = f(\mathbf{k}, x, y, z, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3} v_x dy dz dt. \quad (5.14)$$

За это же время изменение количества электронов $dN_{\text{п}}$ в результате ухода их из объема через правую грань составит:

$$dN_{\text{п}} = -f(\mathbf{k}, x + dx, y, z, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3} v_x dy dz dt. \quad (5.15)$$

За время dt количество электронов в элементе объема $d\tau_r$ изменится на величину

$$\begin{aligned} dN_{\text{л}} + dN_{\text{п}} &= [f(\mathbf{k}, x, y, z, t) - f(\mathbf{k}, x + dx, y, z, t)] \frac{d\tau_{\mathbf{k}}}{4\pi^3} v_x dy dz dt = \\ &= -v_x \frac{\partial f}{\partial x} \frac{\partial \tau_{\mathbf{k}}}{4\pi^3} dx dy dz dt = -v_x \frac{\partial f}{\partial x} \frac{\partial \tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt. \end{aligned} \quad (5.16)$$

В общем случае произвольного направления движения носителей заряда со скоростью $\mathbf{v}$ с составляющими v_x, v_y, v_z изменение числа электронов с данным $\mathbf{k}$ в элементе объема фазового пространства $d\Gamma$ за время dt составит величину, равную

$$- \left[v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y} + v_z \frac{\partial f}{\partial z} \right] \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt = -(\mathbf{v} \nabla_{\mathbf{r}} f) \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt. \quad (5.17)$$

Это изменение количества электронов обусловлено процессами переноса, происходящими в результате диффузии носителей заряда при наличии градиента их концентрации или градиента температуры.

Рассмотрим теперь движение электронов в пространстве волновых векторов. Подсчитаем изменение числа электронов за время dt за счет прихода и ухода их через соответствующие грани элемента объема $d\tau_{\mathbf{k}}$. Для этого нужно, так же как и в случае движения электронов в обычном пространстве, подсчитать количество электронов, пришедших и ушедших через все грани элемента объема $d\tau_{\mathbf{k}}$. В результате получим, что при движении электронов в пространстве волновых векторов их количество в элементе объема фазового пространства $d\Gamma$ за время dt изменится на величину:

$$\begin{aligned} &- \left[\frac{\partial k_x}{\partial t} \frac{\partial f}{\partial k_x} + \frac{\partial k_y}{\partial t} \frac{\partial f}{\partial k_y} + \frac{\partial k_z}{\partial t} \frac{\partial f}{\partial k_z} \right] \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt = \\ &= - \left[\frac{d\mathbf{k}}{dt} \nabla_{\mathbf{k}} f \right] \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt = - \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt. \end{aligned} \quad (5.18)$$

Здесь

$$\mathbf{F}(\mathbf{r}, t) = \frac{d\mathbf{p}}{dt} = \hbar \frac{d\mathbf{k}}{dt} \quad (5.19)$$

есть внешняя сила, действующая на электрон в точке $\mathbf{r}$ в момент времени t .

Соотношение (5.18) выражает изменение количества электронов в результате действия внешней силы, обусловленной электрическим и магнитным полями.

Кроме того, число электронов в элементе объема фазового пространства $d\Gamma$ может меняться в результате взаимодействия электронов с локальными нарушениями периодичности поля кристаллической решетки. Можно считать, что область действия локального возмущения распространяется на расстояние около 10^{-7} см, которое охватывает несколько периодов решетки. Так как тепловая скорость электрона в полупроводнике при комнатной температуре порядка 10^7 см/с, то время взаимодействия электрона со структурными нарушениями кристалла составляет 10^{-14} с. Такое кратковременное взаимодействие не может привести к заметному изменению координаты, но зато сильно изменит скорость и квазимпульс электрона. Этот процесс аналогичен удару в механике, поэтому он и получил название столкновения или соударения. При соударении происходит изменение количества электронов, движущихся в данном направлении, поэтому процессы столкновения называют также процессами рассеяния.

Предположим, что при рассеянии электроны переходят из состояний $(\mathbf{r}, \mathbf{k})$ в состояние $(\mathbf{r}', \mathbf{k}')$. Поскольку при столкновении имеет место кратковременное взаимодействие, которое не вызывает изменения координаты, то вероятность перехода электронов не будет зависеть от $\mathbf{r}$ и $\mathbf{r}'$. Обозначим вероятность такого перехода в единицу времени как $W(\mathbf{k}, \mathbf{k}')$. В элементе объема $d\tau_{\mathbf{k}}$ число занятых состояний составляет величину $f(\mathbf{k}, \mathbf{r}, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3}$, а число свободных

состояний в этом объеме соответственно равно $[1 - f(\mathbf{k}, \mathbf{r}, t)] \frac{d\tau_{\mathbf{k}}}{4\pi^3}$.

Каждое столкновение сопровождается переходом электрона из состояния $\mathbf{k}$ в состояние $\mathbf{k}'$. Но чтобы такой переход состоялся, состояние $\mathbf{k}'$ должно быть свободным. Если $f(\mathbf{k}', \mathbf{r}, t) \frac{d\tau_{\mathbf{k}'}}{4\pi^3}$ — число заня-

тых состояний в объеме элемента $d\tau_{\mathbf{k}'}$, то $[1 - f(\mathbf{k}', \mathbf{r}, t)] \frac{d\tau_{\mathbf{k}'}}{4\pi^3}$ есть число свободных состояний. Поэтому число электронов, способных за время dt в результате рассеяния перейти из состояния $\mathbf{k}$ в состояние $\mathbf{k}'$, приводящее к уменьшению количества электронов в элементе объема фазового пространства $d\Gamma$ за время dt , определяется вероятностью перехода $W(\mathbf{k}, \mathbf{k}')$, числом занятых исходных состояний и числом свободных конечных состояний:

$$- W(\mathbf{k}, \mathbf{k}') f(\mathbf{k}, \mathbf{r}, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3} [1 - f(\mathbf{k}', \mathbf{r}, t)] \frac{d\tau_{\mathbf{k}'}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (5.20)$$

При столкновении носителей заряда с дефектами решетки возможен и обратный переход электронов из состояния k' в состояние k . Пусть вероятность такого перехода в единицу времени есть $W(k', k)$. Тогда число электронов, совершивших за время dt обратный переход, равно:

$$W(k', k) f(k', r, t) \frac{d\tau_{k'}}{4\pi^3} [1 - f(k, r, t)] \frac{d\tau_k}{4\pi^3} d\tau_r dt. \quad (5.21)$$

Поэтому количество электронов в элементе объема фазового пространства $d\Gamma$ в результате рассеяния, сопровождающегося прямыми переходами из k в k' и обратными из k' в k , изменится. Если ради сокращения записи f представить как функцию волнового вектора, то это изменение количества электронов за время dt составит

$$\{W(k', k) f(k') [1 - f(k)] - W(k, k') f(k) [1 - f(k')]\} \frac{d\tau_k}{4\pi^3} \frac{d\tau_{k'}}{4\pi^3} d\tau_r dt. \quad (5.22)$$

Чтобы найти общее количество изменений состояний, занятых электронами в элементе объема фазового пространства за время dt в результате рассеяния, необходимо учесть все возможные k' , т. е. необходимо выражение (5.22) проинтегрировать по объему зоны Бриллюэна V_B :

$$\frac{d\tau_k}{4\pi^3} d\tau_r dt \int_{(V_B)} \{W(k', k) f(k') [1 - f(k)] - W(k, k') f(k) [1 - f(k')]\} \frac{d\tau_{k'}}{4\pi^3}. \quad (5.23)$$

Но изменение количества электронов в элементе объема фазового пространства $d\Gamma$, происходящее за время от t до $t + dt$, составляет

$$f(k, r, t + dt) \frac{d\tau_k}{4\pi^3} d\tau_r - f(k, r, t) \frac{d\tau_k}{4\pi^3} d\tau_r = \frac{\partial f}{\partial t} \frac{\partial \tau_k}{4\pi^3} d\tau_r dt. \quad (5.24)$$

Это изменение количества носителей заряда обусловлено движением электронов за счет диффузии, действия на них внешней силы и наличия процессов рассеяния. Следовательно, выражение (5.24) равно сумме выражений (5.17), (5.18) и (5.23). После сокращения $\frac{d\tau_k}{4\pi^3} d\tau_r dt$ получим:

$$\begin{aligned} \frac{\partial f}{\partial t} = & - (v \nabla_r f) - \frac{1}{\hbar} (F \nabla_k f) + \int_{(V_B)} \{W(k', k) f(k') [1 - f(k)] - \\ & - W(k, k') f(k) [1 - f(k')]\} \frac{d\tau_{k'}}{4\pi^3}. \end{aligned} \quad (5.25)$$

Согласно принципу микроскопической обратимости вероятности прямых и обратных переходов равны, т. е.

$$W(k, k') = W(k', k). \quad (5.26)$$

С учетом (5.26) соотношение (5.25) упрощается и принимает вид:

$$\frac{\partial f}{\partial t} = -(\mathbf{v} \nabla_{\mathbf{r}} f) - \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) + \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (5.27)$$

Уравнение (5.27) принято называть кинетическим уравнением Больцмана. Это уравнение является интегродифференциальным, так как под интегралом стоит искомая функция.

Легко видеть физический смысл уравнения (5.27). Левая часть его отражает изменение функции распределения во времени. Первый и второй члены правой части уравнения показывают изменения функции распределения под действием внешних полей, приводящих к движению электронов в обычном пространстве и в пространстве волновых векторов. Изменение функции распределения электронов по состояниям в результате их рассеяния (столкновения) учтено в уравнении последним членом. Таким образом, на носители заряда с одной стороны действуют внешние силы, стремящиеся упорядочить их движение, а с другой — столкновения, стремящиеся разупорядочить их движение. Поэтому уравнение (5.27) можно записать в виде

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{\text{пол}} + \left(\frac{\partial f}{\partial t} \right)_{\text{ст}}. \quad (5.28)$$

Здесь член

$$\left(\frac{\partial f}{\partial t} \right)_{\text{пол}} = -(\mathbf{v} \nabla_{\mathbf{r}} f) - \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) \quad (5.29)$$

носит название полевого члена уравнения Больцмана, а величину

$$\left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = \int_{(V_B)} \left\{ W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3} \right\} \quad (5.30)$$

называют интегралом столкновений (или соударений).

В случае стационарных процессов, когда величины, описывающие явления переноса, не зависят от времени, функция распределения удовлетворяет равенству

$$\frac{\partial f}{\partial t} = 0 \quad (5.31)$$

и кинетическое уравнение Больцмана (5.28) для стационарного состояния имеет вид:

$$\left(\frac{\partial f}{\partial t} \right)_{\text{пол}} = - \left(\frac{\partial f}{\partial t} \right)_{\text{ст}} \quad (5.32)$$

или

$$(\mathbf{v} \nabla_{\mathbf{r}} f) + \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) = \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (5.33)$$

Из выражений (5.32) и (5.33) следует, что в стационарном состоянии изменения неравновесной функции распределения, вызванные

движением носителей заряда и действием на них внешней силы, компенсируются соударениями носителей заряда с локальными нарушениями периодичности поля решетки.

Необходимо указать, что кинетическое уравнение Больцмана по существу является квазиклассическим, поскольку в функции распределения одновременно задаются координата и волновой вектор. В силу этого оно применимо к процессам, для которых имеет смысл понятие длины свободного пробега. Это означает, что средняя длина свободного пробега носителей заряда должна быть больше длины волны де Бройля. Кроме того, это уравнение можно использовать для описания явления переноса в магнитном поле, если это поле не меняет спектр энергий носителей заряда в кристалле, т. е. не становятся существенными квантовые эффекты. Это условие накладывает ограничение на значение напряженности магнитного поля.

5.3. РАВНОВЕСНОЕ СОСТОЯНИЕ

При термодинамическом равновесии, при котором состояние электронов описывается равновесной функцией распределения (5.10), движение электронов в кристалле происходит под воздействием теплоты, поэтому для равновесного случая левая часть уравнения (5.33) равна нулю. Это значит, что

$$\int_{V_B} W(\mathbf{k}, \mathbf{k}') [f_0(\mathbf{k}') - f_0(\mathbf{k})] d\tau_{\mathbf{k}'} = 0. \quad (5.34)$$

Из этого соотношения следует, что в системе, находящейся в равновесном состоянии, имеет место равенство потоков носителей заряда для прямого и обратного процессов, что согласуется с принципом детального равновесия.

На основании (5.34) получаем:

$$f_0(\mathbf{k}') = f_0(\mathbf{k}). \quad (5.35)$$

Подставляя в (5.35) выражение для равновесной функции распределения (5.10), будем иметь:

$$E_1 - F_1 = E_2 - F_2, \quad (5.36)$$

где E_1 , F_1 и E_2 , F_2 — полная энергия и уровень Ферми для электронов с волновыми векторами $\mathbf{k}$ и $\mathbf{k}'$ соответственно. Так как в условиях термодинамического равновесия полная энергия электрона не меняется, то $E_1 = E_2$, а поэтому из (5.36) следует, что

$$F_1 = F_2. \quad (5.37)$$

Это выражение показывает, что в равновесном состоянии во всех частях системы, между которыми может происходить переход электронов, положение уровня Ферми одинаково.

5.4. ВРЕМЯ РЕЛАКСАЦИИ

Решение кинетического уравнения (5.27) и даже для стационарного случая (5.33) является весьма сложным. Но оно значительно упрощается для тех случаев, когда можно ввести время релаксации. Как оказывается, описание процессов рассеяния при помощи времени релаксации возможно, если столкновения частиц упругие, т. е. такие, при которых энергия носителя заряда мало изменяется, и если процессы рассеяния приводят к случайному распределению носителей заряда по скоростям, т. е. имеет место равновероятное рассеяние носителей заряда по всем направлениям.

Допустим, что в момент времени $t = 0$ на систему, описываемую неравновесной функцией распределения f , перестали действовать внешние возмущения (выключаются поля) и полевой член обращается в нуль. В результате процессов соударений система придет в равновесное состояние, описываемое равновесной функцией распределения f_0 . Это значит, что после выключения внешнего поля функция распределения изменяется благодаря наличию соударений электронов с дефектами решетки:

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{\text{ст}}. \quad (5.38)$$

В том случае, когда отклонение распределения носителей заряда от равновесного состояния невелико, можно положить, что в отсутствие внешних полей скорость изменения функции распределения вследствие соударений пропорциональна величине отклонения функции от равновесия, т. е. пропорциональна $f - f_0$:

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = \frac{f - f_0}{\tau(\mathbf{k})}, \quad (5.39)$$

где $1/\tau(\mathbf{k})$ — коэффициент пропорциональности, зависящий от $\mathbf{k}$. Решая уравнения (5.39), получаем:

$$f - f_0 = (f - f_0)_{t=0} e^{-t/\tau}. \quad (5.40)$$

Из (5.40) следует, что после прекращения действия внешних полей разность $(f - f_0)$ уменьшается по экспоненциальному закону с постоянной времени τ , которая носит название времени релаксации. Следовательно, τ есть среднее время, в течение которого в системе существует неравновесное распределение носителей заряда после снятия внешних полей. Поскольку установление равновесного состояния происходит в результате столкновений и при этом достаточно нескольких соударений, то время релаксации должно быть примерно равным времени свободного пробега электрона. С учетом этого можно считать, что длина свободного пробега l определяется скоростью движения электрона v и временем релаксации $\tau(\mathbf{k})$.

Используя (5.39), интеграл столкновений (5.30) можно представить в виде

$$\left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3} = - \frac{f(\mathbf{k}) - f_0(\mathbf{k})}{\tau(\mathbf{k})}. \quad (5.41)$$

Поскольку внешние силы в соответствии с равенством (5.39) не приводят к большим отклонениям распределения носителей заряда от равновесного, то неравновесную функцию распределения можно представить как

$$f(\mathbf{k}) = f_0 + f_1(\mathbf{k}), \quad (5.42)$$

где $f_1(\mathbf{k})$ — поправка к равновесной функции распределения, появившаяся в результате внешнего воздействия на систему частиц.

Запишем ее в виде

$$f_1(\mathbf{k}) = -\frac{\partial f_0}{\partial E} \mathbf{k} \chi(E), \quad (5.43)$$

где $\chi(E)$ — некоторый вектор, зависящий лишь от энергии носителя заряда.

Будем считать, что время релаксации не зависит от внешних полей. Это значит, что $\tau(\mathbf{k})$ определяет и процессы рассеяния при наличии внешних полей. В силу этого для уравнения (5.33) можно использовать интеграл столкновения (5.41) и кинетическое уравнение для стационарного случая примет вид:

$$(\mathbf{v} \nabla_{\mathbf{r}} f) + \frac{1}{h} (\mathbf{F} \nabla_{\mathbf{k}} f) = -\frac{f - f_0}{\tau(\mathbf{k})} = -\frac{f_1(\mathbf{k})}{\tau(\mathbf{k})}. \quad (5.44)$$

Запишем интеграл столкновения (5.41), используя соотношение (5.42)

$$\left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f_1(\mathbf{k}') - f_1(\mathbf{k})] d\tau_{\mathbf{k}'} = -\frac{f_1(\mathbf{k})}{\tau(\mathbf{k})}. \quad (5.45)$$

Из равенства (5.45) можно определить время релаксации. Если учесть соотношение (5.43), то $\tau(\mathbf{k})$ будет определяться равенством следующего вида:

$$\begin{aligned} \frac{1}{\tau(\mathbf{k})} &= -\frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \frac{f_1(\mathbf{k}') - f_1(\mathbf{k})}{f_1(\mathbf{k})} d\tau_{\mathbf{k}'} = \\ &= \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left[1 - \frac{\frac{\partial f_0}{\partial E} \mathbf{k}' \chi(E')}{\frac{\partial f_0}{\partial E} \mathbf{k} \chi(E)} \right] d\tau_{\mathbf{k}'}. \end{aligned} \quad (5.46)$$

Соударения, как известно, могут быть упругими и неупругими. Соударения носят название упругих, если кинетическая энергия сталкивающихся частиц сохраняется. Неупругими соударениями являются такие, в результате которых энергия частиц меняется и после удара становится больше или меньше, чем до удара.

Будем считать, что соударения электронов со структурными нарушениями являются упругими. В этом случае у электрона изменяется лишь направление волнового вектора и, следовательно,

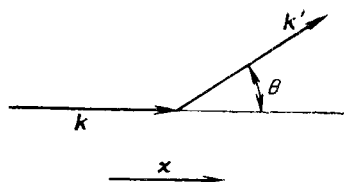


Рис. 5.2. Изменение волнового вектора электрона при рассеянии

направление скорости движения, а их абсолютные значения остаются неизменными. Поэтому при упругом соударении можно считать, что

$$k' = k; \quad v' = v \quad \text{и} \quad E' = E, \quad (5.47)$$

если зоны энергии сферические.

В силу этого время релаксации для упругого соударения с учетом (5.47) будет равно

$$\frac{1}{\tau(k)} = \frac{1}{4\pi^3} \int_{(V_B)} W(k, k') \left[1 - \frac{k' \chi(E)}{k \chi(E)} \right] d\tau_{k'}. \quad (5.48)$$

Пусть векторы χ , k , k' направлены так, как это изображено на рис. 5.2. В этом случае проекции векторов k' и k на вектор χ соответственно равны:

$$k_\chi = k; \quad k'_\chi = k' \cos \theta = k \cos \theta, \quad (5.49)$$

где θ — угол отклонения электрона от первоначального направления движения в результате столкновения. С учетом (5.49) равенство (5.48) запишется в виде

$$\begin{aligned} \frac{1}{\tau(k)} &= \frac{1}{4\pi^3} \int_{(V_B)} W(k, k') \left[1 - \frac{k'_\chi}{k_\chi} \right] d\tau_{k'} = \\ &= \frac{1}{4\pi^3} \int_{(V_B)} W(k, k') (1 - \cos \theta) d\tau_{k'}. \end{aligned} \quad (5.50)$$

Таким образом, при рассмотрении процессов переноса основной задачей является вычисление неравновесной функции распределения, с помощью которой можно рассчитать все величины, характеризующие кинетические эффекты. Но, как следует из (5.44), эта задача решается при условии, если известно время релаксации (5.50), а следовательно, необходимо знание механизма рассеяния носителей заряда.

Рассмотрим теперь рассеяние, приводящее к случайному распределению носителей заряда по скоростям, т. е. к такому, при котором электрон, рассеиваясь, попадает в состояние k' , находящееся возле того же минимума энергии, что и k , но с противоположно направленной скоростью. Такое рассеяние возможно в том случае, если поверхности постоянной энергии являются сферами. Поскольку $W(k, k') = W(k', k)$ — четная функция, а $f_1(k') = -\frac{\partial f_0}{\partial f}(k' \chi)$ — нечетная функция относительно инверсии $k \rightleftharpoons -k$, то

$$\int_{(V_B)} W(k, k') f_1(k') d\tau_{k'} = - \int_{(V_B)} \frac{\partial f_0}{\partial E} W(k, k') (k' \chi) d\tau_{k'} = 0, \quad (5.51)$$

так как нечетная функция при интегрировании в симметричных пределах дает нуль.

С учетом (5.51) интеграл столкновения (5.45) будет равен:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -\frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') f_1(\mathbf{k}) d\tau_{\mathbf{k}'} = -\frac{f_1(\mathbf{k})}{\tau(\mathbf{k})}, \quad (5.52)$$

откуда следует, что время релаксации при рассеянии, приводящем к случайному распределению носителей заряда по скоростям, определяется выражением

$$\frac{1}{\tau(\mathbf{k})} = \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') d\tau_{\mathbf{k}'} \quad (5.53)$$

5.5. РАССЕЯНИЕ НА ИОНАХ ПРИМЕСИ

В примесном полупроводнике каждый ион примеси создает вокруг себя электрическое поле. Под воздействием этого поля движущийся носитель заряда отклоняется от своего первоначального направления (рис. 5.3). При этом носитель заряда отклоняется тем сильнее, чем медленнее и ближе к иону он движется.

Задача сводится фактически к нахождению траектории движения заряженной частицы в центральном поле иона. Она может быть решена чисто классическими методами. Для этого будем считать носитель заряда классической частицей, движущейся в кулоновском поле положительного иона примеси с зарядом Ze . В этом случае потенциальную энергию взаимодействия можно представить в виде

$$U(r) = \pm \frac{Ze^2}{4\pi\epsilon_r\epsilon_0 r}, \quad (5.54)$$

где плюс соответствует дырке, минус — электрону.

Как показывает расчет, в кулоновском поле иона носитель заряда движется по гиперболе. На рис. 5.4 изображены траектории движения электрона и дырки в поле положительного иона. Оба носителя заряда отклоняются одинаково, хотя в одном случае действует сила притяжения, в другом — отталкивания. Через b обозначено прицельное расстояние — минимальное расстояние между ионом и траекторией движения носителя заряда, а θ — угол рассеяния, т. е. угол между первоначальным направлением движения носителя

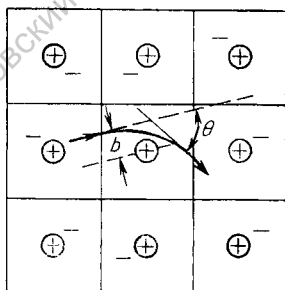


Рис. 5.3. Искривление траектории движения электрона в поле положительного иона примеси

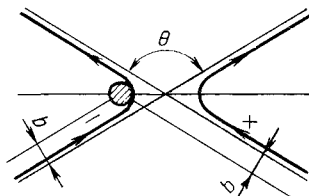


Рис. 5.4. Рассеяние электрона и дырки на положительном ионе примеси

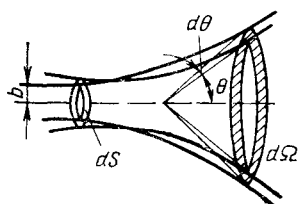


Рис. 5.5. Рассеяние электронов точечным несовершенством

заряда и направлением движения после рассеяния. При этом угол рассеяния зависит от прицельного расстояния следующим образом:

$$b = \frac{Ze^2}{\epsilon_r m^* v^2} \operatorname{ctg} \frac{\theta}{2}, \quad (5.55)$$

где m^* и v — эффективная масса и скорость движения носителя заряда.

Процесс рассеяния является случайным процессом, поэтому различные носители заряда будут отклоняться на различные углы от направления своего движения.

В силу этого при подсчете времени релаксации необходимо учитывать усредненное значение сечения по углам рассеяния.

Допустим, что угол рассеяния меняется от θ до $\theta + d\theta$. Поскольку в нашем случае рассеивающий центр обладает осевой симметрией, то эти углы являются углами двух конусов, как это изображено на рис. 5.5. Здесь через $d\Omega$ обозначен телесный угол, заключенный между этими конусами, он равен:

$$d\Omega = 2\pi \sin \theta d\theta. \quad (5.56)$$

Процессы рассеяния в данном случае характеризуются дифференциальным сечением рассеяния $\sigma(\theta)$. Согласно (5.3) $\sigma(\theta)$ равно отношению числа частиц, отклоненных одним центром в 1 с на угол $d\theta$ в единичный телесный угол, к потоку падающих частиц на 1 см² в 1 с.

Предположим, что n электронов движется беспорядочно со скоростью v . Если выбрать перпендикулярно к каждому направлению движения электронов площадку в 1 см², то на нее за 1 с упадет nv электронов. Полное число частиц, отклоненных за 1 с на угол θ в пределах телесного угла $d\Omega$, будет равно потоку частиц, падающих на кольцо (рис. 5.5) площадью $dS = 2\pi b |db|$. Поэтому на основании определения и равенства (5.56) можно записать:

$$\sigma(\theta) = \frac{nv2\pi b |db|}{nv d\Omega} = \frac{b |db|}{\sin \theta d\theta}. \quad (5.57)$$

Так как из формулы (5.55) имеем:

$$|db| = \frac{Ze^2}{2\epsilon_r m^* v^2} \frac{1}{\sin \frac{\theta}{2}} d\theta, \quad (5.58)$$

то, подставляя выражения (5.55) и (5.58) в (5.57), получаем:

$$\begin{aligned} \sigma(\theta) &= \frac{Ze^2}{\epsilon_r m^* v^2} \operatorname{ctg} \frac{\theta}{2} \frac{Ze^2}{2\epsilon_r m^* v^2} \frac{1}{\sin^2 \frac{\theta}{2}} d\theta \frac{1}{2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} d\theta} = \\ &= \left(\frac{Ze^2}{2\epsilon_r m^* v^2} \right)^2 \frac{1}{\sin^4 \frac{\theta}{2}}. \end{aligned} \quad (5.59)$$

Формула (5.59) была получена Резерфордом при изучении рассеяния α -частиц на ядрах элементов.

Для подсчета времени релаксации необходимо выразить вероятность перехода $W(k, k')$ через дифференциальное сечение рассеяния $\sigma(\theta)$. Предположим, что в кристалле единичного объема имеется N_I ионов примеси. Тогда число рассеянных носителей заряда всеми центрами в 1 с в телесный угол $d\Omega$ составит:

$$W(k, k') n d\Omega = \sigma(\theta) N_I n v d\Omega, \quad (5.60)$$

откуда

$$W(k, k') = \sigma(\theta) N_I v. \quad (5.61)$$

Таким образом, *вероятность перехода из состояния k в состояние k' при рассеянии электронов на ионах примеси определяется дифференциальным эффективным сечением, количеством центров рассеяния и скоростью движения носителей заряда.*

Соударения носителей заряда с ионами примеси можно считать упругими по двум причинам. Во-первых, масса электронов в 10^4 раз меньше массы иона, являющегося рассеивающим центром. Во-вторых, ионы решетки гораздо менее подвижны относительно своего положения равновесия, чем свободные носители заряда, поэтому при соударении носителя заряда с ионом примеси энергия носителя заряда изменяется очень незначительно.

Если при упругом соударении носители заряда рассеиваются в телесный угол $d\Omega$, то время релаксации, используя (5.50), можно записать в виде

$$\frac{1}{\tau(k)} = \int_{(\theta)} W(k, k') (1 - \cos \theta) d\Omega. \quad (5.62)$$

Подставляя формулы (5.56), (5.61) в (5.62), получаем:

$$\frac{1}{\tau(k)} = 2\pi N_I v(k) \int_{(\theta)} \sigma(\theta) (1 - \cos \theta) \sin \theta d\theta = N_I v(k) \sigma_c. \quad (5.63)$$

Величина

$$\sigma_c = 2\pi \int_{(\theta)} \sigma(\theta) (1 - \cos \theta) \sin \theta d\theta \quad (5.64)$$

есть усредненное дифференциальное эффективное сечение рассеяния; усреднение по углам θ производится с весовой функцией $1 - \cos \theta$. σ_c называют эффективным сечением проводимости (или подвижности), а также транспортным эффективным сечением.

Время релаксации на основании (5.63) равно

$$\tau(k) = \frac{1}{N_I \sigma_c v(k)}, \quad (5.65)$$

а длина свободного пробега

$$l(k) = \tau(k) v(k) = \frac{1}{\sigma_c N_I}. \quad (5.66)$$

Из (5.59) следует, что рассеяние носителей заряда на ионах примеси в достаточной степени анизотропно. Рассеяние под малыми углами или при больших значениях прицельного расстояния является преобладающим. Если в формулу (5.64) подставить значение $\sigma(\theta)$ из (5.59), то интеграл на нижнем пределе при $\theta = 0$ будет расходящимся, а сечение рассеяния — бесконечным. Но рассеяние носителей заряда на ионизованной примеси обладает конечным сечением, поэтому при определении времени релаксации интегрирование (5.63) проводится от θ_{min} до π . Определить значение b_{max} , соответствующее θ_{min} , можно, исходя из следующих соображений. Если ионы примеси распределены равномерно по кристаллу и их концентрация равна N_I , то среднее расстояние между ними будет $N_I^{-1/3}$. Сферу действия каждого рассеивающего иона можно ограничить половиной среднего расстояния между ними, поэтому можно считать, что верхний предел прицельного расстояния $b_{max} = \frac{1}{2} N_I^{-1/3}$. Он соответствует наименьшему значению угла рассеяния θ_{min} , который можно определить из (5.55):

$$\operatorname{ctg} \frac{\theta_{min}}{2} = \frac{\epsilon_r m^* v^2}{Ze^2} b_{max} = \frac{\epsilon_r m^* v^2}{2Ze^2 N_I^{1/3}}. \quad (5.67)$$

Учитывая, что $1 - \cos \theta = 2 \sin^2 \frac{\theta}{2}$; $\sin \theta = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}$, а

$$\int_{\theta_{min}}^{\pi} \frac{(1 - \cos \theta) \sin \theta d\theta}{\sin^4 \frac{\theta}{2}} = 8 \int_{\theta_{min}}^{\pi} \frac{\cos \frac{\theta}{2} d\left(\frac{\theta}{2}\right)}{\sin \frac{\theta}{2}} = -4 \ln \sin^2 \frac{\theta_{min}}{2},$$

и воспользовавшись выражениями (5.67) для $\operatorname{ctg} \frac{\theta_{min}}{2}$ и (5.59) для $\sigma(\theta)$, найдем эффективное сечение проводимости:

$$\begin{aligned} \sigma_e &= 2\pi \int_{\theta_{min}}^{\pi} \sigma(\theta) (1 - \cos \theta) \sin \theta d\theta = \\ &= 8\pi \left(\frac{Ze^2}{2\epsilon_r m^* v^2} \right)^2 \ln \frac{1}{\sin^2 \frac{\theta_{min}}{2}} = 2\pi \left(\frac{Ze^2}{\epsilon_r m^* v^2} \right)^2 \ln \left[1 + \operatorname{ctg}^2 \frac{\theta_{min}}{2} \right] = \\ &= 2\pi \left(\frac{Ze^2}{\epsilon_r m^* v^2} \right)^2 \ln \left[1 + \left(\frac{\epsilon_r m^* v^2}{2Ze^2 N_I^{1/3}} \right)^2 \right]. \end{aligned} \quad (5.68)$$

Время релаксации при рассеянии носителей заряда на ионизованных атомах примеси на основании (5.65) равно:

$$\tau_I(k) = \frac{1}{\sigma_e N_I v(k)} = \frac{\epsilon_r^2 m^{*2} v^3(k)}{2\pi Z^2 e^4 N_I \ln \left[1 + \left(\frac{\epsilon_r m^* v^2}{2Ze^2 N_I^{1/3}} \right)^2 \right]}. \quad (5.69)$$

Это соотношение для τ принято называть формулой Конуэлл—Вайскопфа.

Если выразить скорость через энергию $v = (2E/m^*)^{1/2}$, то время релаксации при рассеянии носителей заряда на ионах примеси примет вид:

$$\tau_I = \tau_{0I} E^{3/2}, \quad (5.70)$$

где

$$\tau_{0I} = \frac{(2m^*)^{1/2} e_r^2}{\pi Z^2 e^4 N_I \ln \left[1 + \left(\frac{\epsilon_r E}{Z e^2 N_I^{1/3}} \right)^2 \right]}, \quad (5.71)$$

а длина свободного пробега равна:

$$l_I(k) = \tau_I(k) v(k) = \left(\frac{2}{m^*} \right)^{1/2} \tau_{0I} E^2. \quad (5.72)$$

Таким образом, при рассеянии носителей заряда на ионах примеси время релаксации зависит от энергии в степени $3/2$ и пропорционально эффективной массе в степени $1/2$.

5.6. РАССЕЯНИЕ НА АТОМАХ ПРИМЕСИ И ДИСЛОКАЦИЯХ

Рассеяние носителей заряда на нейтральных атомах примеси проявляется гораздо слабее, чем рассеяние на ионах примеси. Однако при низких температурах, когда концентрация ионизованных атомов примеси меньше концентрации нейтральных атомов примеси, этот механизм рассеяния играет существенную роль. В этом случае рассеяние носителей заряда может осуществляться двумя путями: при помощи прямого упругого соударения и обменом падающего электрона с электроном примесного атома. Определение времени релаксации для этих двух механизмов рассеяния достаточно сложно. Однако на основе расчета, при котором рассеяние носителей заряда на нейтральных атомах примеси рассматривалось как рассеяние медленных электронов на атоме водорода, находящемся в среде с относительной диэлектрической проницаемостью ϵ_r , получена зависимость

$$\tau_A = \frac{e^2 m^{*2}}{20 \epsilon_r \hbar^3} \frac{1}{N_A}, \quad (5.73)$$

где N_A — концентрация нейтральных атомов примеси при данной температуре.

Как следует из (5.73), время релаксации при рассеянии носителей заряда на нейтральной примеси не зависит ни от температуры, ни от энергии рассеиваемого носителя заряда, но определяется концентрацией атомов примеси. Поэтому этот механизм рассеяния особенно важен при очень низких температурах, когда примесь не ионизована.

Рассеяние носителей заряда может происходить на нарушениях решетчатых, вносимых дислокациями. Вокруг дислокации существует область, в которой кристалл деформирован, поэтому в окрестности дислокации зависимость энергии от волнового вектора отличается

от закона дисперсии в идеальном кристалле. Это приводит к рассеянию свободных носителей заряда. В полупроводниках, которые кристаллизуются в структуре типа алмаза, дислокации действуют как цепочки акцепторов, которые захватывают свободные электроны. Поэтому они представляют собой положительный цилиндр с отрицательно заряженной линией в середине, т. е. дислокацию можно уподобить линейному отрицательному заряду в виде бесконечно длинного цилиндра радиусом R , вокруг которого имеется положительный пространственный заряд. Электроны, движущиеся перпендикулярно дислокациям, не могут свободно перемещаться в направлении поля, так как должны обойти дислокации и рассеяться на них, испытывая отталкивание. Расчет показывает, что *при рассеянии носителей заряда на дислокациях время релаксации не зависит от температуры кристалла, а определяется плотностью дислокаций на единицу поверхности N_D и скоростью рассеиваемого электрона v согласно выражению вида*

$$\tau_D = \frac{3}{8Rv} \frac{1}{N_D}. \quad (5.74)$$

Если принять $v = 10^7$ см/с; $R = 3 \cdot 10^{-5}$ см; $N_D = 10^6$ см $^{-2}$, то время релаксации $\tau_D = 1,25 \cdot 10^{-9}$ с. Эта величина на 4—5 порядков больше, чем время релаксации для рассеяния на тепловых колебаниях решетки при комнатной температуре. Поэтому это рассеяние незначительно при комнатных температурах, но оно может быть существенно при низких температурах.

5.7. РАССЕЯНИЕ НА ТЕПЛОВЫХ КОЛЕБАНИЯХ РЕШЕТКИ

Взаимодействие электрона (или дырки) с колебаниями решетки осуществляется двояким образом: 1) электрон передает часть своей энергии решетке, так что определенное нормальное колебание с частотой ω_q увеличивает свое квантовое число на единицу. В результате такого процесса рассеяния образуется фотон с энергией $\hbar\omega_q$ и квазиимпульсом $\mathbf{Q} = \hbar\mathbf{q}$ и число фононов возрастает на единицу; $\mathbf{Q} = \hbar\mathbf{q}$ есть импульс, который приобретает решетка вследствие соударения с электронами; 2) электрон при взаимодействии получает часть энергии от решетки, так что квантовое число определенного нормального колебания с частотой ω_q уменьшается на единицу. В этом случае исчезает фотон с энергией $\hbar\omega_q$ и квазиимпульсом $\mathbf{Q} = \hbar\mathbf{q}$, а число фононов уменьшается на единицу. Следовательно, в любом из этих двух случаев электрон, сталкиваясь с фононом, обменивается с ним энергией и квазиимпульсом. Такой механизм рассеяния называют *однофононным*. Так как число фононов определяется температурой, то и рассеяние электронов на тепловых колебаниях решетки должно зависеть от температуры. В кристалле возможен и многофононный механизм рассеяния, при котором имеет место поглощение или испускание более чем одного фонона. Однако вероятность такого процесса мала по сравнению с вероятностью однофононного рассеяния.

Допустим, что до столкновения электрон имел волновой вектор $\mathbf{k}$ и энергию $E(\mathbf{k})$, а после столкновения — волновой вектор $\mathbf{k}'$ и энергию $E(\mathbf{k}')$. Фонон, возникающий или исчезающий при столкновении, обладает энергией $\hbar\omega_q$ и волновым вектором $\mathbf{q}$. При взаимодействии электрона с фононом должны выполняться два закона: закон сохранения энергии и закон сохранения квазиимпульса. Поэтому при поглощении электроном фонона

$$E'(\mathbf{k}') = E(\mathbf{k}) + \hbar\omega_q; \quad (5.75)$$

$$\mathbf{k}' = \mathbf{k} + \mathbf{q}, \quad (5.76)$$

а число фононов, обладающих волновым числом q , в конечном состоянии уменьшится на единицу, т. е.

$$N'_q = N_q - 1. \quad (5.77)$$

При испускании электроном фонона

$$E(\mathbf{k}') = E(\mathbf{k}) - \hbar\omega_q; \quad (5.78)$$

$$\mathbf{k}' = \mathbf{k} - \mathbf{q}, \quad (5.79)$$

при этом

$$N'_q = N_q + 1. \quad (5.80)$$

Рассеяние носителя заряда может быть упругим и неупругим. Для характеристики рассеяния необходимо знание относительного изменения энергии δ электрона за одно столкновение

$$\delta = \frac{E' - E}{E} = \frac{\Delta E}{E} \quad (5.81)$$

или за единицу времени

$$A = \frac{\Delta E}{\tau} = \frac{E\delta}{\tau}. \quad (5.82)$$

В случае изотропного и упругого рассеяния $\delta \ll 1$ и среднее время рассеяния τ имеет смысл времени релаксации системы по импульсу, а $\tau_E = \tau/\delta$ — времени релаксации по энергии.

Рассмотрим процесс рассеяния электронов в атомном полупроводнике кубической структуры, энергетическая схема которого представлена на рис. 5.6, где a — постоянная решетки в отсутствие деформации. Будем считать, что зона проводимости такого кристалла простая, так что энергия электрона в зоне проводимости равна

$$E = \frac{\hbar^2 k^2}{2m^*}. \quad (5.83)$$

В твердом теле поперечные волны представляют собой волны деформации сдвига и в кубическом кристалле не приводят к изменению объема. Продольные волны, являющиеся волнами деформации сжатия и растяжения, вызывают изменение объема кристалла. Как следует из рис. 5.6, при сжатии кристалла, сопровождающемся

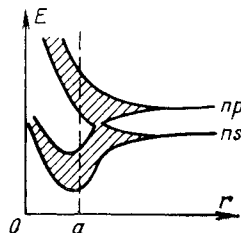


Рис. 5.6. Энергетическая схема атомного полупроводника



Рис. 5.7. Изменение энергии дна зоны проводимости и потолка валентной зоны атомного полупроводника под воздействием продольных акустических колебаний решетки

увеличением постоянной решетки a , нижний край зоны проводимости смещается вверх, а верхний край валентной зоны — вниз, в результате этого ширина запрещенной зоны увеличится. При растяжении, приводящем к увеличению постоянной решетки a , ширина запрещенной зоны уменьшится. Следовательно, в таком кристалле локальная деформация, создаваемая продольной акустической волной, приводит к волнообразному смещению дна зоны проводимости и потолка валентной зоны, как это схематически представлено на рис. 5.7. Поэтому движущийся электрон, сталкиваясь с волной смещения, обусловленной тепловыми колебаниями решетки, будет рассеиваться только на продольных колебаниях.

Покажем, что это рассеяние упругое и происходит на длинноволновых акустических фононах. Закон сохранения энергии согласно (5.75), (5.79) и (5.83) запишется в виде

$$\frac{\hbar^2 k'^2}{2m^*} = \frac{\hbar^2 (k \pm q)^2}{2m^*} = \frac{\hbar^2 k^2}{2m^*} \pm \hbar \omega_q, \quad (5.84)$$

откуда

$$q^2 \pm 2kq \cos \vartheta \mp \frac{2m^*}{\hbar} \omega_q = 0, \quad (5.85)$$

где ϑ — угол между направлениями векторов $\mathbf{k}$ и $\mathbf{q}$.

В равенстве (5.84) знак плюс относится к случаю поглощения, а минус — к случаю испускания фонона. Если электрон взаимодействует только с длинноволновыми акустическими фононами, то $\omega_q = v_{зв} q$, где $v_{зв}$ — скорость распространения продольной звуковой волны. Учитывая это, из (5.85) находим:

$$q = \mp 2k \cos \vartheta \pm \frac{2m^* v_{зв}}{\hbar}. \quad (5.86)$$

Так как $\frac{\hbar^2 k^2}{2m^*} = \langle E \rangle = kT$, то $k = \frac{1}{\hbar} \sqrt{2m^* kT}$ и

$$\frac{m^* v_{зв}}{\hbar k} = \sqrt{\frac{m^* v_{зв}^2}{2kT}} = \sqrt{\frac{T_{кр}}{T}}, \quad (5.87)$$

где $T_{кр} = \frac{m^* v_{зв}^2}{2k}$.

Если принять $v_{зв} = 3 \cdot 10^5$ см/с, $m^* = m_0 = 10^{-27}$ г, а $k = 1,38 \cdot 10^{-23}$ Дж·К, то $T_{кр} = 1$ К. Поэтому при температурах, намного превышающих 1 К, можно пренебречь вторым слагаемым в (5.86) по сравнению с первым и записать:

$$q = \mp 2k \cos \vartheta. \quad (5.88)$$

В зависимости от ϑ волновое число q может меняться от $q_{min} = 0$ до $q_{max} = 2k$, т. е. в среднем на k . При $T = 300$ К $k = 10^7$ см $^{-1}$

и q может принимать значения от 0 до $2 \cdot 10^7 \text{ см}^{-1}$, что соответствует изменению энергии фононов от 0 до $\hbar\omega_q = \hbar v_{\text{зв}} q = \hbar v_{\text{зв}} k = 2 \cdot 10^{-3} \text{ эВ}$. Но при комнатной температуре энергия электронов $E = \hbar^2 k^2 / 2m^* = 2,5 \cdot 10^{-2} \text{ эВ}$, поэтому

$$\hbar\omega_q/E \approx 8 \cdot 10^{-2} \ll 1. \quad (5.89)$$

Максимальное значение волнового вектора q_0 по теории Дебая согласно (3.87) и (3.96) равно: $q_0 = (6\pi^2 N/V)^{1/3} = (6\pi^2 N/a^3)^{1/3} \approx 10^8 \text{ см}^{-1}$, следовательно, $q_{\text{max}} \ll q_0$. Отсюда следует, что *рассеяние электронов на продольных акустических колебаниях является упругим; при этом электрон рассеивается на длинноволновых фононах, поглощая или излучая фонон с $q = k$* . Так как энергия фононов значительно меньше энергии электронов, то $k' = k$ и $E' = E$, поскольку в (5.84) можно не учитывать энергию фонона.

При низких температурах средняя скорость движения электрона приближается к скорости распространения упругих волн в кристалле и энергией фонона в (5.84) пренебречь нельзя. В этом случае имеет место неупругое столкновение электрона с фононом.

Поскольку электроны рассеиваются на продольных акустических фононах упруго, то для определения времени релаксации можно воспользоваться формулой (5.48):

$$\begin{aligned} \frac{1}{\tau} &= \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left[1 - \frac{k'_x}{k_x} \right] d\tau_{\mathbf{k}'} = \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left(\mp \frac{q_x}{k_x} \right) d\tau_{\mathbf{k}'} = \\ &= - \frac{1}{4\pi^3} \int_{(V_B)} W^+(\mathbf{k}, \mathbf{k}') \frac{q_x}{k_x} d\tau_{\mathbf{k}'} + \frac{1}{4\pi^3} \int_{(V_B)} W^-(\mathbf{k}, \mathbf{k}') \frac{q_x}{k_x} d\tau_{\mathbf{k}'} \quad (5.90) \end{aligned}$$

Здесь первый член правой части равенства учитывает поглощение, а второй — испускание фононов при рассеянии электронов.

Вероятность перехода в 1 с электрона из состояния $\mathbf{k}$ в состояние $\mathbf{k}'$, как показывает расчет *, равна в случае поглощения фонона:

$$W^+(\mathbf{k}, \mathbf{k}') = W^+(\mathbf{k}, \mathbf{k} + \mathbf{q}) = \omega(q) N_q \delta(E' - E - \hbar\omega_q), \quad (5.91)$$

а в случае испускания фонона

$$W^-(\mathbf{k}, \mathbf{k}') = W^-(\mathbf{k}, \mathbf{k} - \mathbf{q}) = \omega(q) (N_q + 1) \delta(E' - E + \hbar\omega_q), \quad (5.92)$$

где $\delta(E' - E \pm \hbar\omega_q)$ — дельта-функция, она выражает законы сохранения энергии и волнового вектора;

$$\omega(q) = \frac{4\pi}{9N} \frac{C_2 q^2}{M\omega_q}; \quad (5.93)$$

C — постоянная, имеющая размерность энергии, характеризует интенсивность взаимодействия электронов с колебаниями решетки, N — число элементарных ячеек в кристалле; M — масса атома.

* Ансельм А. И. Введение в теорию полупроводников. М., Наука, 1978, с. 471.

В электронных процессах переноса функция распределения фононов N_q мало отличается от равновесной $\langle n \rangle$ (3.75), поэтому можно считать, что

$$N_q = \langle n \rangle = \frac{1}{e^{\hbar\omega_q/kT} - 1}. \quad (5.94)$$

Согласно (5.87) $\frac{\hbar\omega_q}{kT} = \frac{\hbar v_{\text{зв}q}}{kT} = \frac{\hbar v_{\text{зв}k}}{kT} = \sqrt{\frac{2T_{\text{кр}}}{T}} \ll 1$, поэтому выражение (5.94) можно разложить в ряд, в результате будем иметь:

$$\langle n \rangle = \langle n + 1 \rangle = \frac{kT}{\hbar v_{\text{зв}q}}. \quad (5.95)$$

Если в (5.90) подставить (5.91) и (5.92) с учетом (5.94) и (5.95) и провести интегрирование по q в пределах от $q_{\text{min}} = 0$ до $q_{\text{max}} = 2k$, получим:

$$\tau_l = \frac{9\pi}{4\sqrt{2}} \frac{\hbar^4 v_{\text{зв}}^2 M}{C^2 a^3 k m^{3/2}} T^{-1} E^{-1/2} = \frac{\tau_{0l}}{m^{3/2}} T^{-1} E^{-1/2}, \quad (5.96)$$

где $\tau_{0l} = 9\pi \hbar^4 v_{\text{зв}}^2 M / 4 \sqrt{2} a^3 k C^2$; a — параметр решетки.

Длина свободного пробега

$$l = \tau_l v = \sqrt{2} \tau_{0l} T^{-1} / m^{3/2}. \quad (5.97)$$

Таким образом, время релаксации при рассеянии электронов на длинноволновых акустических колебаниях решетки в атомных полупроводниках обратно пропорционально температуре в первой степени и энергии в степени $1/2$, а длина свободного пробега не зависит от энергии носителей заряда.

Оценим время релаксации для германия в случае рассеяния электронов на колебаниях решетки. В чистом германии при комнатной температуре подвижность электронов $\mu_n = 0,39 \text{ м}^2/(\text{В} \cdot \text{с})$. Если принять $m_n^* = 0,3m_0$, то $e/m_n^* = 6 \cdot 10^{11} \text{ Кл/кг}$, следовательно, $\tau_l = m_n^* \mu_n / e = 6,5 \cdot 10^{-13} \text{ с}$.

Из сравнения соотношений (5.96) и (5.70) следует, что при повышении энергии носителей заряда время релаксации при рассеянии их на ионизованной примеси увеличивается, а при рассеянии на акустических колебаниях уменьшается. Это означает, что в случае примесного рассеяния преобладающее влияние оказывают медленные электроны и дырки, а поэтому этот механизм рассеяния должен сильнее проявляться при низких температурах. При рассеянии на акустических фононах основной вклад вносят сравнительно быстрые электроны и этот механизм рассеяния должен быть значительнее при высоких температурах.

В полярных полупроводниках, а также в полупроводниковых соединениях типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$, в которых связь между атомами носит частично ионный характер, электроны проводимости гораздо сильнее взаимодействуют с оптическими колебаниями, чем с акустическими. Поскольку при оптических колебаниях смещение ионов в ячейке происходит в противоположные стороны, то разноимен-

ные заряды создадут электрические поля, перемещающиеся по кристаллу в виде плоских волн. Длинноволновые оптические колебания получили название *поляризационных волн*. Взаимодействие носителей заряда с поляризационными волнами и приводит к их рассеянию. При этом продольные колебания сильнее рассеивают, чем поперечные. Каждое столкновение электрона с оптическим фононом приводит к возникновению или исчезновению фонона с энергией $\hbar\omega_{\text{оп}}$. Частота оптического фонона слабо зависит от квазимпульса q , поэтому при соударении энергия электрона либо увеличивается, либо уменьшается на одинаковую величину, равную:

$$\pm \hbar\omega_{\text{оп}} = \pm \hbar\omega_0, \quad (5.98)$$

где ω_0 — максимальная частота продольной оптической ветви колебания.

Энергия, передаваемая носителем заряда решетке за одно столкновение, определяется энергией возникшего фонона, умноженной на отношение разности вероятностей испускания и поглощения к сумме этих вероятностей. Но вероятность поглощения фонона согласно (5.91) пропорциональна концентрации фононов N_q , а вероятность испускания фонона согласно (5.92) пропорциональна $N_q + 1$. Следовательно, энергия, передаваемая электроном фонону, равна:

$$\begin{aligned} \Delta E &= \hbar\omega_0 \frac{W^-(\mathbf{k}, \mathbf{k}') - W^+(\mathbf{k}, \mathbf{k}')}{W^-(\mathbf{k}, \mathbf{k}') + W^+(\mathbf{k}, \mathbf{k}')} = \hbar\omega_0 \frac{(N_q + 1) - N_q}{(N_q + 1) + N_q} = \\ &= \hbar\omega_0 \frac{1}{2N_q + 1} = \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2kT}. \end{aligned} \quad (5.99)$$

Из формулы (5.99) следует, что на одно столкновение электрона, сопровождающееся передачей энергии фонону, приходится $2N_q + 1$ столкновений.

При высоких температурах, когда $T \gg \theta_0$ или $kT \gg \hbar\omega_0$, т. е. когда энергия оптического фонона $\hbar\omega_0$ много меньше энергии электрона kT , тогда

$$\delta_{\text{оп}} = \frac{\Delta E}{E} = \frac{(\hbar\omega_0)^2}{2(kT)^2} \ll 1 \quad (5.100)$$

и рассеяние можно считать упругим.

В этом случае у большинства электронов изменение энергии при поглощении или испускании оптического фонона будет незначительно. Поэтому можно ввести время релаксации τ и определить его по формуле (5.90). Для случая рассеяния электронов на оптических фононах в ионном кристалле при высоких температурах, как показывает расчет, время релаксации

$$\tau_{\text{оп}} = \frac{V\sqrt{2} \hbar^2 e_r}{2e^2 k m^{*1/2}} T^{-1} E^{1/2}, \quad (5.101)$$

а длина свободного пробега

$$l_{\text{оп}} = \frac{\hbar^2 e_r}{e^2 m^* k} T^{-1} E. \quad (5.102)$$

При низких температурах, когда $T \ll \theta_0$ или $kT \ll \hbar\omega_0$, электроны могут только поглощать оптические фононы, в результате чего энергия их сильно увеличивается. Поскольку вероятность испускания фонона (5.92) в $(N_q + 1)/N_q$ раз превышает вероятность поглощения фонона (5.91), время пребывания электрона в состоянии с большой энергией будет очень мало. Можно считать, что электрон сразу после поглощения фонона испускает фотон той же энергии. Фактически рассеяние можно рассматривать как процесс обмена фононами, а поэтому энергия электрона практически остается неизменной, но направление его квазиимпульса сильно изменится. Это позволяет считать, что и при низких температурах, значительно меньших характеристической температуры $\theta_0 = \hbar\omega_0/k$, электрон, взаимодействуя с оптическими колебаниями решетки, рассеивается упруго.

Если учесть, что в случае низких температур

$$N_q = \langle n \rangle = \frac{1}{e^{\hbar\omega_0/kT} - 1} \approx e^{-\hbar\omega_0/kT}, \quad (5.103)$$

расчет времени релаксации по (5.90), в которой учитывается только поглощение фонона, дает

$$\tau_{\text{оп}} = \frac{3\sqrt{2}\hbar^2\varepsilon_r e^{\hbar\omega_0/kT}}{2e^2 m^{*1/2} (\hbar\omega_0)^{1/2}}, \quad (5.104)$$

которое не зависит от энергии электрона, но экспоненциально зависит от температуры.

Длина свободного пробега электронов в случае рассеяния на оптических фононах при низких температурах определяется соотношением вида

$$l_{\text{оп}} = \frac{3\hbar^2\varepsilon_r e^{\hbar\omega_0/kT}}{e^2 m^{*3/2} (\hbar\omega_0)^{1/2}} E^{1/2}, \quad (5.105)$$

т. е. зависит от энергии носителей заряда в степени $1/2$.

ГЛАВА ШЕСТАЯ

КИНЕТИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

6.1. НЕРАВНОВЕСНАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

При описании явлений переноса с помощью кинетического уравнения Больцмана необходимо знание неравновесной функции распределения для электронов и дырок. Ее определение проведем для невырожденного полупроводника, находящегося в электрическом и магнитном полях и имеющего градиент температуры. Для полупроводника со сферическими изоэнергетическими поверхностями энергия электронов

$$E = \frac{\hbar^2 k^2}{2m_n^*}, \quad (6.1)$$

а их скорость

$$v = \frac{e}{\hbar} \frac{\partial E}{\partial \mathbf{k}} = \frac{\hbar}{m_n^*} \mathbf{k}. \quad (6.2)$$

При наличии внешних полей — электрического напряженностью $\mathcal{E}$ и магнитного с индукцией $\mathbf{B}$ — на электрон действует сила

$$\mathbf{F} = -e (\mathcal{E} + [\mathbf{v}\mathbf{B}]) \quad (6.3)$$

и кинетическое уравнение для стационарного случая (5.44) примет вид:

$$(\mathbf{v}\nabla_{\mathbf{r}}f) - \frac{e}{\hbar} (\mathcal{E} + [\mathbf{v}\mathbf{B}]) \nabla_{\mathbf{k}}f = -f_1(\mathbf{k})/\tau_e, \quad (6.4)$$

где τ_e — время релаксации электрона; $\mathcal{E} = -\nabla\varphi$; φ — электростатический потенциал.

Будем считать, что неравновесная функция распределения $f(\mathbf{k}) = f_0 + f_1(\mathbf{k})$ мало отличается от равновесной f_0 . Тогда в уравнении (6.4) можно заменить f на f_0 везде, кроме члена, содержащего магнитное поле. Для этого случая уравнение (6.4) запишется в виде

$$(\mathbf{v}\nabla_{\mathbf{r}}f_0) - \frac{e}{\hbar} (\mathcal{E}\nabla_{\mathbf{k}}f_0) - \frac{e}{\hbar} [\mathbf{v}\mathbf{B}] \nabla_{\mathbf{k}}f = -\frac{f_1(\mathbf{k})}{\tau_e}. \quad (6.5)$$

Для невырожденного электронного газа

$$\frac{\partial f_0}{\partial E} = -\frac{1}{kT} e^{-(E-F)/kT}; \quad (6.6)$$

$$\frac{\partial f_0}{\partial T} = \frac{E-F}{kT^2} e^{-(E-F)/kT} = \frac{F-E}{T} \frac{\partial f_0}{\partial E}; \quad (6.7)$$

$$\frac{\partial f_0}{\partial F} = \frac{1}{kT} e^{-(E-F)/kT} = -\frac{\partial f_0}{\partial E}, \quad (6.8)$$

поэтому

$$\nabla_{\mathbf{r}}f_0 = \frac{\partial f_0}{\partial T} \nabla T + \frac{\partial f_0}{\partial F} \nabla F = \frac{\partial f_0}{\partial E} \left(\frac{F-E}{T} \nabla T - \nabla F \right), \quad (6.9)$$

а

$$\nabla_{\mathbf{k}}f_0 = \frac{\partial f_0}{\partial E} \frac{\hbar^2}{m_n^*} \mathbf{k} = \hbar \frac{\partial f_0}{\partial E} \mathbf{v} \quad (6.10)$$

и

$$[\mathbf{v}\mathbf{B}] \nabla_{\mathbf{k}}f = [\mathbf{v}\mathbf{B}] \nabla_{\mathbf{k}}(f_0 + f_1) = \hbar \frac{\partial f_0}{\partial E} [\mathbf{v}\mathbf{B}] \mathbf{v} + [\mathbf{v}\mathbf{B}] \nabla_{\mathbf{k}}f_1. \quad (6.11)$$

Но $[\mathbf{v}\mathbf{B}] \mathbf{v} = [\mathbf{v}\mathbf{v}] \mathbf{B} = 0$, а поскольку

$$f_1(\mathbf{k}) = -\frac{\partial f_0}{\partial E} \chi_n \mathbf{k} = -\frac{m_n^*}{\hbar} \frac{\partial f_0}{\partial E} (\chi_n \mathbf{v}), \quad (6.12)$$

то

$$\nabla_{\mathbf{k}}f_1 = -\nabla_{\mathbf{k}} \left\{ \frac{\partial f_0}{\partial E} (\chi_n \mathbf{k}) \right\} = -\frac{\partial f_0}{\partial E} \chi_n - \left\{ \hbar \mathbf{k} \frac{\partial}{\partial E} \left(\frac{\partial f_0}{\partial E} \chi_n \right) \right\} \mathbf{v}. \quad (6.13)$$

Тогда

$$(\mathbf{v}\mathbf{B}) \nabla_{\mathbf{k}}f = [\mathbf{v}\mathbf{B}] \nabla_{\mathbf{k}}f_1 = -\frac{\partial f_0}{\partial E} ([\mathbf{v}\mathbf{B}] \chi_n) = -\frac{\partial f_0}{\partial E} [\mathbf{B}\chi_n] \mathbf{v}, \quad (6.14)$$

так как $[\mathbf{v}\mathbf{B}] \chi_n = [\mathbf{B}\chi_n] \mathbf{v}$.

Далее, используя (6.9), (6.8), (6.10) и (6.14), из (6.5) находим:

$$\begin{aligned} f_1(\mathbf{k}) &= -\frac{m_n^*}{\hbar} \frac{\partial f_0}{\partial E} (\chi_n \mathbf{v}) = \\ &= -\tau_e \frac{\partial f_0}{\partial E} \left\{ \frac{F-E}{T} \nabla T - \nabla (F - e\varphi) + \frac{e}{\hbar} [\mathbf{B} \chi_n] \right\} \mathbf{v}. \end{aligned} \quad (6.15)$$

Следовательно, на основании (6.12)

$$\chi_n = -\frac{\hbar}{m_n^*} \tau_e \left\{ \frac{E-F}{T} \nabla T + \nabla (F - e\varphi) - \frac{e}{\hbar} [\mathbf{B} \chi_n] \right\}. \quad (6.16)$$

Поскольку энергия E' и уровень Ферми F' дырок равны:

$$E' = -\frac{\hbar^2 \mathbf{k}'^2}{2m_p^*}; \quad (6.17)$$

$$F' = -E_g - F, \quad (6.18)$$

где E_g — ширина запрещенной зоны, то

$$\chi_p = -\frac{\hbar}{m_p^*} \tau_h \left\{ \frac{E' + E_g + F}{T} \nabla T - \nabla (F - e\varphi) + \frac{e}{\hbar} [\mathbf{B} \chi_p] \right\}. \quad (6.19)$$

Здесь τ_h — время релаксации дырки.

Определение χ_n и χ_p из (6.16) и (6.19) сводится к решению векторного уравнения вида

$$\mathbf{x} = \mathbf{a} + [\mathbf{b}\mathbf{x}]. \quad (6.20)$$

Умножим уравнение (6.20) скалярно на вектор $\mathbf{b}$:

$$(\mathbf{b}\mathbf{x}) = (\mathbf{b}\mathbf{a}) + \mathbf{b}[\mathbf{b}\mathbf{x}] = (\mathbf{b}\mathbf{a}). \quad (6.21)$$

Умножим теперь уравнение (6.20) векторно на вектор $\mathbf{b}$:

$$[\mathbf{b}\mathbf{x}] = [\mathbf{b}\mathbf{a}] + [\mathbf{b}[\mathbf{b}\mathbf{x}]] = [\mathbf{b}\mathbf{a}] + \mathbf{b}(\mathbf{b}\mathbf{x}) - \mathbf{x}(\mathbf{b}\mathbf{b}) = [\mathbf{b}\mathbf{a}] + \mathbf{b}(\mathbf{b}\mathbf{a}) - \mathbf{x}b^2. \quad (6.22)$$

Тогда

$$\mathbf{x} = \mathbf{a} + [\mathbf{b}\mathbf{a}] + \mathbf{b}(\mathbf{b}\mathbf{a}) - \mathbf{x}b^2, \quad (6.23)$$

откуда

$$\mathbf{x} = \frac{\mathbf{a} + [\mathbf{b}\mathbf{a}] + \mathbf{b}(\mathbf{b}\mathbf{a})}{1 + b^2}. \quad (6.24)$$

Если воспользоваться (6.24), то выражения (6.16) и (6.19) запишутся в виде

$$\begin{aligned} \chi_n = -\frac{\hbar \tau_e}{m_n^*} & \left[\frac{\frac{E-F}{T} \nabla T + \nabla (F - e\varphi) + \frac{e\tau_e}{m_n^*} \left[\mathbf{B} \left\{ \frac{E-F}{T} \nabla T + \nabla (F - e\varphi) \right\} \right]}{1 + \left(\frac{e\tau_e}{m_n^*} B \right)^2} + \right. \\ & \left. + \frac{\left(\frac{e\tau_e}{m_n^*} \right)^2 \left(\mathbf{B} \left\{ \frac{E-F}{T} \nabla T + \nabla (F - e\varphi) \right\} \right) \mathbf{B}}{1 + \left(\frac{e\tau_e}{m_n^*} B \right)^2} \right]; \end{aligned} \quad (6.25)$$

$$\chi_p = -\frac{\hbar\tau_h}{m_p^*} \times$$

$$\times \left[\frac{\frac{E' + E_g + F}{T} \nabla T - \nabla (F - e\varphi) - \frac{e\tau_h}{m_p^*} \left[\mathbf{B} \left\{ \frac{E' + E_g + F}{T} \nabla T - \nabla (F - e\varphi) \right\} \right]}{1 + \left(\frac{e\tau_h}{m_p^*} B \right)^2} + \right.$$

$$\left. + \frac{\left(\frac{e\tau_h}{m_p^*} \right)^2 \left(\left\{ \mathbf{B} \frac{E' + E_g + F}{T} \nabla T - \nabla (F - e\varphi) \right\} \right) \mathbf{B}}{1 + \left(\frac{e\tau_h}{m_p^*} B \right)^2} \right]. \quad (6.26)$$

6.2. УДЕЛЬНАЯ ЭЛЕКТРИЧЕСКАЯ ПРОВОДИМОСТЬ ПОЛУПРОВОДНИКОВ

Рассчитаем удельную электрическую проводимость однородного невырожденного полупроводника при отсутствии градиента температуры ($\nabla T = 0$) и магнитного поля ($\mathbf{B} = 0$), изоэнергетические поверхности которого представляют собой сферы. Пусть проводник находится в постоянном однородном электрическом поле напряженностью $\mathcal{E}$.

В элементе объема $d\tau_k$ кристалла единичного объема согласно (5.12) количество электронов

$$dn = f(\mathbf{k}) \frac{d\tau_k}{4\pi^3}. \quad (6.27)$$

Эти электроны, движущиеся под действием внешнего электрического поля со скоростью $\mathbf{v}$, создают элементарную плотность тока

$$d\mathbf{J} = -e\mathbf{v}f(\mathbf{k}) \frac{d\tau_k}{4\pi^3}. \quad (6.28)$$

Если в полупроводнике имеется два сорта носителей заряда, то полная плотность тока

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p = -\frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_n f_n(\mathbf{k}) d\tau_k + \frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_p f_p(\mathbf{k}') d\tau_{k'}, \quad (6.29)$$

где $\mathbf{v}_n$, $f_n(\mathbf{k})$ и $\mathbf{v}_p$, $f_p(\mathbf{k}')$ — скорость и неравновесная функция распределения электронов и дырок соответственно. Интегрирование нужно проводить по всей зоне Бриллюэна.

Если в (6.29) поставить равновесную функцию распределения f_0 , то

$$\mathbf{J} = \int_{(V_B)} \mathbf{v} f_0(\mathbf{k}) d\tau_k = 0, \quad (6.30)$$

$f_0(\mathbf{k})$ — четная функция, а $\mathbf{v}_0(\mathbf{k}) = \frac{\hbar}{m^*} \mathbf{k} f_0(\mathbf{k})$ — нечетная функция и дает нуль при интегрировании в симметричных пределах.

Выражение (6.30) означает, что при термодинамическом равновесии в веществе нет тока.

Если, например, неравновесная функция

$$f_n(\mathbf{k}) = f_0 + f_1^{(n)}(\mathbf{k}) = f_0 - \frac{\partial f_0}{\partial E}(\chi_n \mathbf{k}), \quad (6.31)$$

то с учетом (6.30) вместо выражения (6.29) будем иметь:

$$\begin{aligned} \mathbf{J} &= -\frac{e}{4\pi^3} \int_{(V_B)} v_n f_1^{(n)}(\mathbf{k}) d\tau_k + \frac{e}{4\pi^3} \int_{(V_B)} v_p f_1^{(p)}(\mathbf{k}') d\tau_{k'} = \\ &= \frac{e\hbar}{4\pi^3 m_n^*} \int_{(V_B)} \frac{\partial f_0}{\partial E}(\chi_n \mathbf{k}) \mathbf{k} d\tau_k - \frac{e\hbar}{4\pi^3 m_p^*} \int_{(V_B)} \frac{\partial f_0}{\partial E'}(\chi_p \mathbf{k}') \mathbf{k}' d\tau_{k'}. \end{aligned} \quad (6.32)$$

Поскольку полупроводник однородный, то $\nabla_{\mathbf{r}} f = 0$ и при отсутствии градиента температуры ($\nabla T = 0$) на основании равенства (6.9)

$$\nabla F = 0. \quad (6.33)$$

При наличии электрического поля $\mathcal{E} = -\nabla\varphi$ и отсутствии магнитного поля ($\mathbf{B} = 0$), учитывая (6.33), из (6.25) и (6.26) получаем:

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \mathcal{E}; \quad (6.34)$$

$$\chi_p = \frac{e\hbar\tau_h}{m_p^*} \mathcal{E}. \quad (6.35)$$

Используя (6.34), плотность электронного тока запишем в виде

$$\mathbf{J}_n = -\frac{e^2 \hbar^2}{4\pi^3 m_n^{*2}} \int_{(V_B)} \tau_e \frac{\partial f_0}{\partial E}(\mathbf{k}\mathcal{E}) \mathbf{k} d\tau_k. \quad (6.36)$$

Для невырожденного полупроводника при отсчете энергии от дна зоны проводимости количество электронов в зоне проводимости в соответствии с (4.66) определяется соотношением

$$n = N_c e^{-(E_c - F)/kT} = N_c e^{F/kT}. \quad (6.37)$$

Поскольку для невырожденного состояния $f_0 = e^{-(E-F)/kT}$, то

$$\frac{\partial f_0}{\partial E} = -\frac{1}{kT} e^{-(E-F)/kT} = -\frac{n}{kTN_c} e^{-E/kT}. \quad (6.38)$$

Подставляя (6.38) в (6.36), получаем:

$$\mathbf{J}_n = \frac{e^2 \hbar^2 n}{4\pi^3 m_n^{*2} kTN_c} \int_{(V_B)} \tau_e e^{-E/kT} (\mathbf{k}\mathcal{E}) \mathbf{k} d\tau_k. \quad (6.39)$$

Если воспользоваться соотношением

$$\int F(\mathbf{k}) (\mathbf{a}\mathbf{k}) \mathbf{k} d\tau_k = \frac{4\pi}{3} \mathbf{a} \int_4^\infty F(\mathbf{k}) k^4 dk, \quad (6.40)$$

то (6.39) запишется в виде

$$J_n = \frac{e^2 \hbar^2 n \mathcal{E}}{4\pi^3 m_n^* k T N_c} \frac{4\pi}{3} \int_0^\infty \tau_e e^{-E/kT} k^4 dk, \quad (6.41)$$

Полагая $\alpha = E/kT$, будем иметь:

$$k^4 dk = \frac{(2m_n^* kT)^{5/2}}{2\hbar^5} \alpha^{3/2} d\alpha, \quad (6.42)$$

а подставляя (6.42) в (6.41) и учитывая, что $N_c = 2 (2\pi m_n^* kT/\hbar^2)^{3/2}$, получаем:

$$J_n = \frac{e^2 n}{m_n^*} \mathcal{E} \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha. \quad (6.43)$$

Введем обозначение

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha. \quad (6.44)$$

Тогда

$$J_n = \frac{e^2 n \langle \tau_e \rangle}{m_n^*} \mathcal{E} = en\mu_n \mathcal{E} = \sigma_n \mathcal{E}, \quad (6.45)$$

где

$$\mu_n = \frac{e \langle \tau_e \rangle}{m_n^*} \quad (6.46)$$

— дрейфовая подвижность, а

$$\sigma_n = en\mu_n \quad (6.47)$$

— удельная проводимость, обусловленная электронами. Соответственно дырочная составляющая тока

$$J_p = \frac{e^2 p \langle \tau_h \rangle}{m_p^*} \mathcal{E} = ep\mu_p \mathcal{E} = \sigma_p \mathcal{E}. \quad (6.48)$$

Здесь среднее время релаксации дырок

$$\langle \tau_h \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_h e^{-\alpha} \alpha^{3/2} d\alpha, \quad (6.49)$$

дрейфовая подвижность дырок

$$\mu_p = \frac{e \langle \tau_h \rangle}{m_p^*} \quad (6.50)$$

и удельная проводимость, обусловленная дырками,

$$\sigma_p = ep\mu_p. \quad (6.51)$$

Следовательно, полная плотность тока в полупроводнике, имеющем два типа носителей заряда, равна:

$$J = J_n + J_p = (\sigma_n + \sigma_p) \mathcal{E} = \sigma \mathcal{E}, \quad (6.52)$$

где $\sigma = \sigma_n + \sigma_p$ — удельная проводимость полупроводника.

6.3. ЗАВИСИМОСТЬ ПОДВИЖНОСТИ НОСИТЕЛЕЙ ЗАРЯДА ОТ ТЕМПЕРАТУРЫ

Как следует из (6.46) и (6.50), для вычисления подвижности носителей заряда необходимо знать среднее время релаксации. Однако на время релаксации существенным образом влияет механизм рассеяния, поэтому при столкновении носителей заряда с различными по природе дефектами кристаллической решетки подвижность по-разному будет зависеть от температуры.

А. Рассеяния на колебаниях решетки

В атомных полупроводниках при рассеянии носителей заряда на акустических колебаниях решетки время релаксации согласно (5.96) равно:

$$\tau_l = \frac{\tau_{0l}}{m^{*3/2}} T^{-1} E^{-1/2}. \quad (6.53)$$

На основании (6.44) найдем:

$$\langle \tau_l \rangle = \frac{4}{3\sqrt{\pi}} \frac{\tau_{0l}}{m^{*3/2} k^{1/2}} T^{-3/2} \quad (6.54)$$

и

$$\mu_l = \frac{4e}{3\sqrt{\pi}} \frac{\tau_{0l}}{k^{1/2} m^{*6/2}} T^{-3/2}. \quad (6.55)$$

Подвижность можно выразить через длину свободного пробега. Используя формулу (5.97), получаем:

$$\mu_l = \frac{4el}{3(2\pi m^* kT)^{1/2}}. \quad (6.56)$$

Из (6.55) следует, что в атомных полупроводниках подвижность носителей заряда при рассеянии их на тепловых колебаниях решетки уменьшается с ростом температуры. Кроме того, подвижность обратно пропорциональна эффективной массе носителей заряда в степени 5/2. Поскольку $m_n^* < m_p^*$, то подвижность электронов больше подвижности дырок, а у полупроводников с малой эффективной массой носителей заряда подвижность очень большая (см. табл. 6.1).

Б. Рассеяние на ионизованных примесях

При рассеянии на ионах примеси время релаксации в зависимости от энергии согласно (5.70) и (5.71)

$$\tau_l = \frac{\sqrt{2} \epsilon_r^2 m^{*1/2} E^{3/2}}{\pi Z^2 e^4 N_l \ln [1 + (\epsilon_r E / Ze^2 N_l^{1/3})^2]}. \quad (6.57)$$

При усреднении этого выражения логарифмический член, являющийся медленно меняющейся функцией, можно вынести за знак интеграла, если в этот член подставить такое значение энергии, при котором оставшаяся часть подынтегрального выражения достигает максимума. Для этого необходимо положить $E = 3kT$. По-

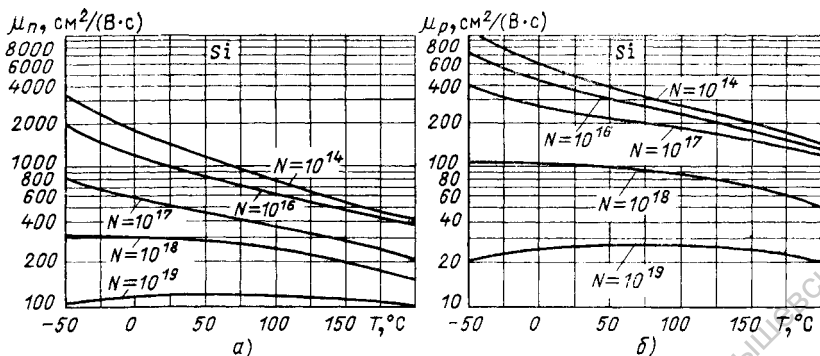


Рис. 6.1. Зависимость дрейфовой подвижности электронов (а) и дырок (б) от температуры в образцах кремния, имеющих разную концентрацию примеси (см⁻³)

этому среднее время релаксации носителей заряда при рассеянии на ионах примеси будет равно:

$$\langle \tau_l \rangle = \frac{8 \sqrt{2} \epsilon_r^2 k^{3/2} m^{1/2} T^{3/2}}{\pi^{3/2} Z^2 e^4 N_l \ln [1 + (3 \epsilon_r k T / Z e^2 N_l^{1/3})^2]}, \quad (6.58)$$

а подвижность при рассеянии ионами примеси запишется в виде

$$\mu_l = \frac{8 \sqrt{2} \epsilon_r^2 k^{3/2} T^{3/2}}{\pi^{3/2} Z^2 e^3 N_l m^{1/2} \ln [1 + (3 \epsilon_r k T / Z e^2 N_l^{1/3})^2]}. \quad (6.59)$$

При достаточно больших температурах можно считать, что

$$\mu_l \approx \mu_{0l} T^{3/2}. \quad (6.60)$$

Из формулы (6.60) следует, что подвижность носителей заряда, обусловленная рассеянием на ионах примеси, уменьшается при уменьшении температуры. Это происходит потому, что с увеличением температуры кристалла возрастает тепловая скорость носителей заряда, а это приводит к ослаблению взаимодействия носителей заряда с ионизованными атомами примеси, так как уменьшается длительность взаимодействия. Кроме того, при данной температуре подвижность уменьшается с увеличением концентрации примеси, которая усиливает эффект рассеяния. Это согласуется с экспериментальными данными. В качестве примера на рис. 6.1 представлена зависимость подвижности электронов и дырок от температуры в образцах кремния, имеющих разную концентрацию примесей.

В. Рассеяние на нейтральных примесях

При рассеянии носителей заряда на нейтральной примеси время релаксации согласно выражению (5.73) не зависит от температуры и энергии носителя заряда. В силу этого подвижность носителей заряда, обусловленная рассеянием на нейтральных атомах примеси, в явном виде не зависит от температуры и будет равна

$$\mu_A = e^2 m^* / 20 \epsilon_r \hbar^3 N_A. \quad (6.61)$$

Г. Рассеяние на дислокациях

При рассеянии носителей заряда на дислокациях подвижность

$$\mu_D = \frac{eT^{-1/2}}{2^{3/2}\pi^{1/2}k^{1/2}m^{*1/2}RN_D}. \quad (6.62)$$

Д. Сложный механизм рассеяния

В реальных полупроводниках строгая периодичность поля решетки нарушена в результате наличия в ней разных дефектов. В силу этого механизм рассеяния носит сложный характер. Если считать, что все механизмы рассеяния независимы, то полная вероятность рассеяния равна сумме вероятностей рассеяния на каждом типе рассеивающих центров, т. е.

$$W(\mathbf{k}, \mathbf{k}') = \sum_i W_i(\mathbf{k}, \mathbf{k}'). \quad (6.63)$$

В этом случае полное время релаксации τ согласно (5.49) будет равно:

$$\tau^{-1} = \frac{1}{4\pi^3} \int_{(V_B)} \sum_i W_i(\mathbf{k}, \mathbf{k}') (1 - \cos \theta) d\mathbf{k}' = \sum_i \tau_i^{-1} \quad (6.64)$$

или

$$\tau = \left(\sum_i \tau_i^{-1} \right)^{-1}. \quad (6.65)$$

Естественно, что вклад различных механизмов рассеяния в суммарное время релаксации зависит от температуры кристалла. По мере изменения энергии носителей заряда роль одних механизмов рассеяния уменьшается, а роль других возрастает, поэтому характер движения носителей заряда под воздействием внешнего электрического поля зависит от того, какой из механизмов рассеяния является доминирующим в данном интервале температур. При очень низких температурах в выражении (6.65) можно учитывать только рассеяние на атомах примеси и дислокациях. С повышением температуры роль этих механизмов уменьшается по сравнению с рассеянием на ионах примеси. При высоких температурах доминирующим становится рассеяние на фононах. Поэтому можно считать, что комбинированная подвижность определяется рассеянием на ионах примеси и на фононах и согласно (6.65) равна $\mu = (\mu_i^{-1} + \mu_f^{-1})^{-1}$, а в зависимости от температуры с учетом (6.55) и (6.59) будет определяться соотношением вида

$$\frac{1}{\mu} = aT^{-3/2} + bT^{3/2}, \quad (6.66)$$

где a и b — постоянные величины.

Как следует из формулы (6.66), в атомных полупроводниках с ростом температуры подвижность носителей заряда растет пропорционально $T^{3/2}$, если рассеяние происходит только на ионах примеси, затем она проходит через максимум и уменьшается пропорционально $T^{-3/2}$, если рассеивающими центрами являются только длинноволновые акустические фононы.

Т а б л и ц а 6.1. Экспериментальные данные по подвижности электронов и дырок некоторых полупроводников

Полупроводник	Подвижность электронов, $\text{см}^2/(\text{В} \cdot \text{с})$		Подвижность дырок, $\text{см}^2/(\text{В} \cdot \text{с})$		Показатель степени температурной зависимости решеточной подвижности p		$b = \mu_n/\mu_p$	
	300 К	77 К	300 К	77 К	Электроны	Дырки	300 К	77 К
Ge	3900	37 100	1900	43 700	1,66	2,33	2	0,8
Si	1500	45 500	600	11 600	2,5	2,7	2,5	3,9
InSb	78 000	1 200 000	750	10 000	1,6	2,1	100	120
InAs	33 000	820 000	460	690	1,2	2,3	70	120
InP	4600	24 000	150	1200	2,0	2,4	30	20
GaSb	4000	6000	1400	3600	2,0	0,9	3	1,7
GaAs	8500	21 000	420	4200	1,0	2,1	20	5
GaP	110	500	75	420	1,5	1,5	1,5	1,2
AlSb	200	—	420	—	—	1,8	0,2	—

Как показывает эксперимент, у полупроводников действительно наблюдается такой ход зависимости подвижности от температуры. В качестве примера на рис. 6.2 представлены температурные зависимости подвижностей электронов и дырок в антимониде индия.

Но, как следует из данных табл. 6.1, у большинства полупроводников температурная зависимость подвижности при рассеянии на тепловых колебаниях решетки не соответствует теоретическому значению показателя степени $3/2$. Это объясняется тем, что в реальных полупроводниках имеют место и другие механизмы рассеяния, такие как рассеяние на оптических колебаниях, двухфононное рассеяние, рассеяние на носителях заряда.

Е. Температурная зависимость удельной проводимости

Учитывая зависимость подвижности и концентрации носителей заряда от температуры, удельную проводимость собственного полупроводника можно записать в виде

$$\sigma = en_i\mu_n + ep_i\mu_p = AT^{3/2+p}e^{-E_g/2kT}. \quad (6.67)$$

Множитель $T^{3/2+p}$ медленно меняется с температурой, тогда как множитель $e^{-E_g/kT}$ сильно зависит от температуры, если $E_g \gg kT$. Поэтому для не слишком высоких температур можно считать, что

$$AT^{3/2+p} = \sigma_0 = \text{const}, \quad (6.68)$$

и выражение (6.67) можно заменить более простым

$$\sigma = \sigma_0 e^{-E_g/2kT}. \quad (6.69)$$

В примесном полупроводнике, как следует из рис. 6.3, б, при достаточно высоких температурах электропроводность полупроводника является собственной. При низких температурах в области истощения примеси концентрация основных носителей заряда остается постоянной (рис. 6.3, а) и электропроводность меняется вследствие изменения подвижности носителей заряда с температурой.

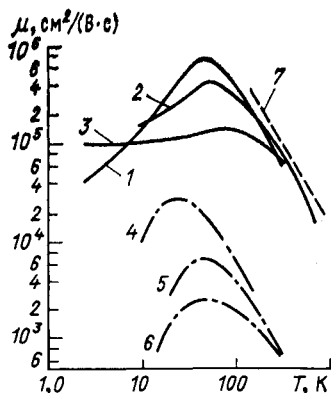
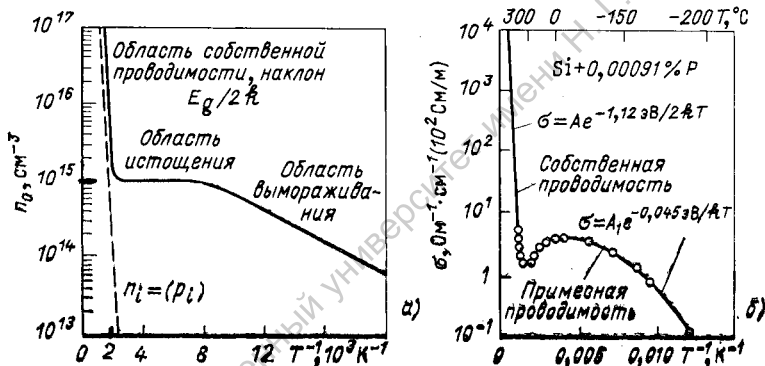


Рис. 6.2. Температурные зависимости подвижности электронов и дырок в селениде индия:

1 — $n = 2 \cdot 10^{14}$; 2 — $n = 10^{15}$; 3 — $n = 10^{16} \text{ см}^{-3}$; 4 — $p = 3 \cdot 10^{14}$; 5 — $p = 4 \cdot 10^{15}$; 6 — $p = 3 \times 10^{16} \text{ см}^{-3}$; 7 — теоретические значения подвижности электронов

Рис. 6.3. Зависимость концентрации носителей заряда (а) и удельной проводимости (б) от обратной температуры для образца кремния с концентрацией донорных примесей около 10^{15} см^{-3}



6.4. ЭФФЕКТ ХОЛЛА

Кинетические эффекты, имеющие место при одновременном воздействии на проводник электрического и магнитного полей, называют гальваномагнитными. Эффект Холла является одним из гальваномагнитных эффектов. Рассмотрение эффекта Холла проведем для слабого магнитного поля. Под слабым магнитным полем понимают такое поле, для которого период обращения T_c по круговой орбите в магнитном поле много больше времени релаксации τ :

$$\tau \ll T_c. \quad (6.70)$$

Но частота обращения носителя заряда с эффективной массой m^* по круговой орбите в магнитном поле с индукцией B , т. е. частота ω_c циклотронного резонанса согласно (2.150), равна:

$$\omega_c = \frac{2\pi}{T_c} = \frac{e}{m^*} B, \quad (6.71)$$

поэтому в слабых полях

$$\frac{\tau\omega_c}{2\pi} = \frac{e\tau}{2\pi m^*} B = \frac{\mu B}{2\pi} \ll 1, \quad (6.72)$$

в сильных полях

$$\frac{\tau\omega_c}{2\pi} = \frac{\mu B}{2\pi} \gg 1. \quad (6.73)$$

Таким образом, условие слабого и сильного магнитного полей определяется не только индукцией магнитного поля B , но и подвижностью носителей заряда μ .

Если воспользоваться длиной свободного пробега $l = v\tau$ и радиусом циклотронной орбиты $r = v/\omega_c$, по которой движется носитель заряда, то

$$l/r = \tau\omega_c. \quad (6.74)$$

Как видно, в слабом магнитном поле, поскольку $r \gg l$, носитель заряда, движущийся по круговому пути в плоскости, перпендикулярной $\mathbf{B}$, успевает пройти до столкновения малое расстояние по круговой орбите, а в сильном магнитном поле, для которого $r \ll l$, траектория носителя заряда искривляется очень сильно.

Допустим, что по полупроводнику, имеющему форму прямоугольной пластинки (рис. 6.4), под действием электрического поля $\mathcal{E}$ протекает ток с плотностью

$$\mathbf{J} = -env = \sigma \mathcal{E}. \quad (6.75)$$

Если полупроводник однородный, то эквипотенциальные поверхности расположены перпендикулярно направлению электрического поля $\mathcal{E}$, а следовательно, и вектору плотности тока $\mathbf{J}$. Поэтому разность потенциалов между точками А и Б на рис. 6.4, лежащими в плоскости, перпендикулярной $\mathbf{J}$, будет равна нулю.

Теперь поместим полупроводник в магнитное поле, перпендикулярное вектору тока, как это показано на рис. 6.4. В этом случае на носитель заряда, движущийся с дрейфовой скоростью $\mathbf{v}$, будет действовать сила Лоренца

$$\mathbf{F} = \pm e[\mathbf{vB}], \quad (6.76)$$

направленная перпендикулярно $\mathbf{v}$ и $\mathbf{B}$. Здесь плюс соответствует дырке, а минус — электрону.

Но

$$\mathbf{v} = \pm \mu \mathcal{E} = \pm \frac{e\langle \tau \rangle}{m^*} \mathcal{E}, \quad (6.77)$$

поэтому

$$\mathbf{F} = \frac{e^2 \langle \tau \rangle}{m^*} [\mathcal{E} \mathbf{B}]. \quad (6.78)$$

Из (6.78) следует, что направление силы Лоренца не зависит от знака носителя заряда, а определяется направлением векторов $\mathcal{E}$ и $\mathbf{B}$ или $\mathbf{J}$ и $\mathbf{B}$. Следовательно, если скорость носителей заряда определяется внешним электрическим полем, то электроны и дырки под действием силы Лоренца отклоняются в одну и ту же сторону.

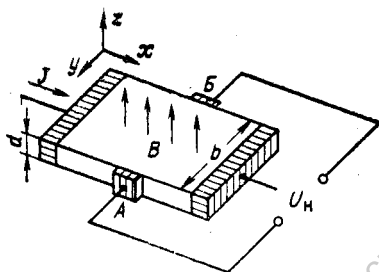


Рис. 6.4. Образец для измерения эффекта Холла

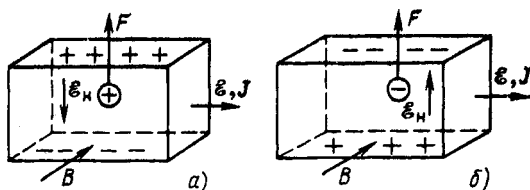


Рис. 6.5. Отклонение носителей заряда под действием магнитного поля в образцах с дырочной (а) и электронной (б) электропроводностью

Для выбранных направлений $\mathbf{B}$ и $\mathcal{E}$, представленных на рис. 6.5, сила Лоренца $\mathbf{F}$ направлена вверх. Под действием этой силы дырки в акцепторном полупроводнике (рис. 6.5, а) и электроны в донорном полупроводнике (рис. 6.5, б) будут отнесены к верхней поверхности образца, вследствие чего на нижней поверхности образца возникнет их дефицит, что обусловит противоположный по знаку заряд по отношению к заряду на верхней поверхности. В результате разделения зарядов появится электрическое поле напряженностью $\mathcal{E}_H$, перпендикулярное направлению магнитного поля. Направление этого поля, которое называют полем Холла, зависит от знака носителей заряда. В нашем случае поле Холла $\mathcal{E}_H$ направлено вниз в p -образце и вверх в n -образце. *Явление возникновения в полупроводнике с текущим по нему током поперечного электрического поля под действием магнитного поля называют эффектом Холла.* Напряженность поля $\mathcal{E}_H$ будет расти до тех пор, пока сила, обусловленная этим полем, не скомпенсирует силу Лоренца:

$$-e\mathcal{E}_H = evB. \quad (6.79)$$

При этом условии носители заряда движутся вдоль образца под действием только продольного электрического поля $\mathcal{E}$, а следовательно, плотность тока $\mathbf{J}$ по направлению совпадает с напряженностью $\mathcal{E}$. Вектор напряженности суммарного электрического поля $\mathcal{E}' = \mathcal{E} + \mathcal{E}_H$, как следует из рис. 6.6, повернут на некоторый угол φ относительно направления тока $\mathbf{J}$. *Угол, заключенный между $\mathbf{J}$ и $\mathcal{E}'$, носит название угла Холла.* Теперь эквипотенциальные поверхности будут повернуты на угол φ относительно первоначального положения, поэтому на образце между точками А и В (рис. 6.4) появится разность потенциалов, называемая ЭДС Холла. Если ширина образцов b , то холловская разность потенциалов

$$V_H = \mathcal{E}_H b = -vBb. \quad (6.80)$$

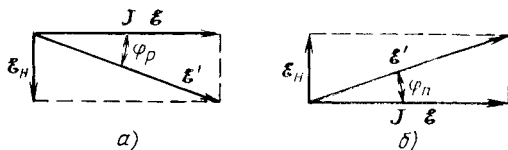
Если воспользоваться (6.75), будем иметь:

$$V_H = -\frac{1}{en} JBb = RJBb. \quad (6.81)$$

Величину R в (6.81) принято называть коэффициентом (или постоянной) Холла, который в случае электронов равен:

$$R = -\frac{1}{en}. \quad (6.82)$$

Рис. 6.6. Угол Холла при дырочной (а) и электронной (б) электропроводности



Если носителями заряда являются дырки, концентрация которых равна p , то, как следует из рис. 6.5, а, е $\mathcal{E}_H = evB$, поэтому

$$R = \frac{1}{ep}. \quad (6.83)$$

Как следует из равенств (6.82) и (6.83), коэффициент Холла обратно пропорционален концентрации носителей заряда, а знак его совпадает со знаком носителей заряда. При таком рассмотрении эффекта Холла не принималось во внимание статистическое распределение носителей заряда по энергиям и не учитывалась зависимость времени релаксации от энергии. Это будет сделано в § 6.5.

Угол Холла φ (рис. 6.6) можно определить из соотношения

$$\operatorname{tg} \varphi = \frac{\mathcal{E}_H}{\mathcal{E}}, \quad (6.84)$$

причем знак $\operatorname{tg} \varphi$ определяется направлением поля Холла: положителен для дырочного полупроводника и отрицателен для электронного. Если магнитное поле слабое, то можно считать, что

$$\varphi = \frac{\mathcal{E}_H}{\mathcal{E}} = \frac{RJB}{\mathcal{E}} = R\sigma B, \quad (6.85)$$

откуда следует, что

$$R = \frac{\varphi}{\sigma B}. \quad (6.86)$$

Учитывая, что для электронного полупроводника

$$\sigma_n = en\mu_n, \quad (6.87)$$

то угол Холла:

$$\varphi_n = R_n \sigma_n B = \mu_n B. \quad (6.88)$$

Аналогично для дырочного полупроводника

$$\varphi_p = R_p \sigma_p B = \mu_p B. \quad (6.89)$$

6.5. ЭФФЕКТ ХОЛЛА В ПОЛУПРОВОДНИКАХ С ДВУМЯ ТИПАМИ НОСИТЕЛЕЙ ЗАРЯДА

Допустим, что в полупроводнике ток переносится электронами и дырками, концентрации и подвижности которых соответственно равны n , μ_n и p , μ_p .

Определение коэффициента Холла проведем для случая слабого магнитного поля. Предположим, что вектор напряженности электрического поля $\mathcal{E}$, вектор плотности тока $\mathbf{J}$ и вектор магнитной индукции $\mathbf{B}$ ориентированы относительно осей координат, как показано на рис. 6.7.

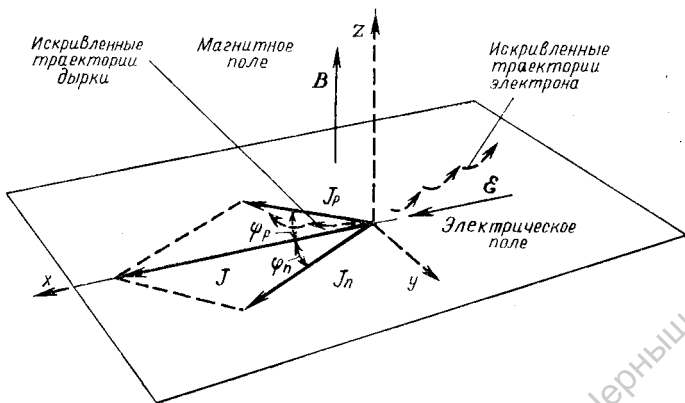


Рис. 6.7. Эффект Холла при наличии двух типов носителей заряда

Под действием электрического поля носители заряда будут ускоряться. По мере увеличения дрейфовой скорости на носитель заряда будет действовать отклоняющая сила Лоренца. Траектория носителя заряда под действием магнитного поля искривляется до тех пор, пока он не испытывает столкновения с дефектами кристаллической решетки (магнитное поле слабое и $l \ll r$). При столкновении носитель заряда передаст решетке энергию, накопленную в электрическом поле, а затем опять будет двигаться под действием поля $\mathcal{E}$. В результате плотность тока дырок $\mathbf{J}_p$ отклоняется от направления электрического поля на угол φ_p . Электроны под действием магнитного поля отклоняются в ту же сторону, что и дырки, а плотность электронного тока $\mathbf{J}_n$ отклоняется от направления электрического поля в противоположную сторону на угол φ_n (рис. 6.7).

При этом суммарная плотность тока, если использовать соотношение (6.32), будет равна:

$$\mathbf{J} = \frac{e\hbar}{4\pi^3 m_n^*} \int_{(V_B)} \frac{\partial f_0}{\partial E} (\chi_n \mathbf{k}) \mathbf{k} d\tau_k - \frac{e\hbar}{4\pi^3 m_p^*} \int_{(V_B)} \frac{\partial f_{0p}}{\partial E'} (\chi_p \mathbf{k}') \mathbf{k}' d\tau_{k'}. \quad (6.90)$$

Вычисление плотности электронного $\mathbf{J}_n$ и дырочного $\mathbf{J}_p$ токов проведем для однородного полупроводника со сферическими изоэнергетическими поверхностями при отсутствии градиента температуры, находящегося в слабом магнитном поле. Так как $\nabla_{\mathbf{r}} f = \frac{E-F}{T} \nabla T + \nabla F = 0$, а $\mathcal{E} = -\nabla \varphi$, то для слабого магнитного поля, когда $\frac{e\tau}{m^*} B = \mu B \ll 1$, (6.25) и (6.26) примут вид:

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B} \mathcal{E}] \right\}; \quad (6.91)$$

$$\chi_p = \frac{e\hbar\tau_h}{m_p^*} \left\{ \mathcal{E} - \frac{e\tau_h}{m_p^*} [\mathbf{B} \mathcal{E}] \right\}. \quad (6.92)$$

Находим теперь электронную составляющую тока, используя (6.91):

$$\mathbf{J}_n = - \frac{e^2 \hbar^2}{4\pi^3 m_n^{*2}} \int_{(V_B)} \frac{\partial f_0}{\partial E} \left\{ \tau_e(\mathbf{k}\boldsymbol{\xi}) + \frac{e\tau_e^2}{m_n^*} \mathbf{k} [\mathbf{B}\boldsymbol{\xi}] \right\} \mathbf{k} d\tau_{\mathbf{k}}. \quad (6.93)$$

Если учесть (6.40), (6.41), (6.42), то уравнение (6.93) запишется в виде

$$\mathbf{J}_n = \frac{e^2 n}{m_n^*} \boldsymbol{\xi} \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha + \frac{e^3 n}{m_n^{*2}} [\mathbf{B}\boldsymbol{\xi}] \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-\alpha} \alpha^{3/2} d\alpha. \quad (6.94)$$

Вводя среднее значение времени релаксации для электронов

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha \quad (6.95)$$

и среднее значение квадрата времени релаксации

$$\langle \tau_e^2 \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-\alpha} \alpha^{3/2} d\alpha, \quad (6.96)$$

а также учитывая, что $\mu_n = e \langle \tau_e \rangle / m_n^*$, получаем для плотности электронной составляющей тока:

$$\mathbf{J}_n = en\mu_n \boldsymbol{\xi} + \frac{\langle \tau_e^2 \rangle}{|\langle \tau_e \rangle|^2} en\mu_n^2 [\mathbf{B}\boldsymbol{\xi}] = en\mu_n \boldsymbol{\xi} + r_n en\mu_n^2 [\mathbf{B}\boldsymbol{\xi}], \quad (6.97)$$

где

$$r_n = \frac{\langle \tau_e^2 \rangle}{|\langle \tau_e \rangle|^2}. \quad (6.98)$$

По аналогии плотность дырочной составляющей тока запишется в виде

$$\mathbf{J}_p = ep\mu_p \boldsymbol{\xi} - r_p ep\mu_p^2 [\mathbf{B}\boldsymbol{\xi}], \quad (6.99)$$

где

$$r_p = \frac{\langle \tau_h^2 \rangle}{|\langle \tau_h \rangle|^2}. \quad (6.100)$$

Полная плотность тока

$$\mathbf{J} = e(n\mu_n + p\mu_p) \boldsymbol{\xi} - r_n e \left(\frac{r_p}{r_n} p\mu_p^2 - n\mu_n^2 \right) [\mathbf{B}\boldsymbol{\xi}]. \quad (6.101)$$

В (6.101) коэффициенты r_n и r_p зависят от механизма рассеяния носителей заряда, который определяет время релаксации $\tau(\mathbf{k})$. В случае упругого рассеяния, когда $\mathbf{k}' = \mathbf{k}$ и $E' = E$, для любого механизма рассеяния отношение

$$\frac{r_p}{r_n} = \frac{\langle \tau_h^2 \rangle}{|\langle \tau_h \rangle|^2} \frac{|\langle \tau_e \rangle|^2}{\langle \tau_e^2 \rangle} = 1. \quad (6.102)$$

Обозначая $r_n = r$ и учитывая (6.102), будем иметь:

$$\mathbf{J} = e(n\mu_n + p\mu_p)\mathbf{E} - re(p\mu_p^2 - n\mu_n^2)[\mathbf{B}\mathbf{E}]. \quad (6.103)$$

Допустим, что плотность тока направлена вдоль оси x , т. е. $J_x = J$ и $J_y = J_z = 0$, а магнитное поле вдоль оси z , т. е. $B_z = B$ и $B_y = B_x = 0$ (рис. 6.7), при этом равенство (6.103) распадается на два уравнения:

$$J_x = e(n\mu_n + p\mu_p)\mathcal{E}_x + re(p\mu_p^2 - n\mu_n^2)B\mathcal{E}_y = J; \quad (6.104)$$

$$J_y = e(n\mu_n + p\mu_p)\mathcal{E}_y - re(p\mu_p^2 - n\mu_n^2)B\mathcal{E}_x = 0. \quad (6.105)$$

Решая эти уравнения относительно $\mathcal{E}_y$, в случае слабого магнитного поля получаем для напряженности поля Холла:

$$\mathcal{E}_y = \frac{r}{e} \frac{p\mu_p^2 - n\mu_n^2}{(n\mu_n + p\mu_p)^2} JB = RJB, \quad (6.106)$$

где коэффициент Холла для полупроводника с двумя типами носителей заряда

$$R = \frac{r}{e} \frac{p\mu_p^2 - n\mu_n^2}{(n\mu_n + p\mu_p)^2}. \quad (6.107)$$

Для собственного полупроводника, у которого $n = p = n_i$,

$$R_i = \frac{r}{en_i} \frac{\mu_p - \mu_n}{\mu_p + \mu_n} = \frac{r}{en_i} \frac{1 - b}{1 + b}. \quad (6.108)$$

Так как обычно $b = \mu_n/\mu_p > 1$, то $R_i < 0$.

Как видим, коэффициент Холла зависит от механизма рассеяния, что учитывается множителем r , который носит название Холл-фактора. Определим r для разных механизмов рассеяния. При рассеянии на акустических колебаниях решетки время релаксации (5.96) с учетом $E = kT\alpha$ равно:

$$\tau_i = \frac{\tau_{0i}}{m_n^{*3/2}} T^{-1} E^{-1/2} = \frac{\tau_{0i}}{m_n^{*3/2} k^{1/2} T^{3/2}} \alpha^{-1/2}; \quad (6.109)$$

на основании (6.95), (6.96) и (6.98) найдем:

$$r = 3\pi/8. \quad (6.110)$$

При рассеянии на ионах примеси, используя соотношение (5.70), получаем:

$$r = 315\pi/512 = 1,93. \quad (6.111)$$

Поскольку рассеяние носителей заряда зависит от температуры, то при определении коэффициента Холла при низких температурах, когда основную роль в рассеянии играют ионы примеси, необходимо полагать $r = 1,93$. Для температур, при которых имеет место рассеяние на акустических фононах, $r = 3\pi/8$. В том случае, когда в процессе рассеяния одновременно участвуют как колебания решетки, так и ионы примеси, r имеет более сложное выражение. На рис. 6.8 показана зависимость r от отношения удельного сопротивления полупроводника ρ_i при рассеянии только на ионах при-

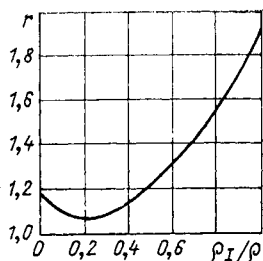
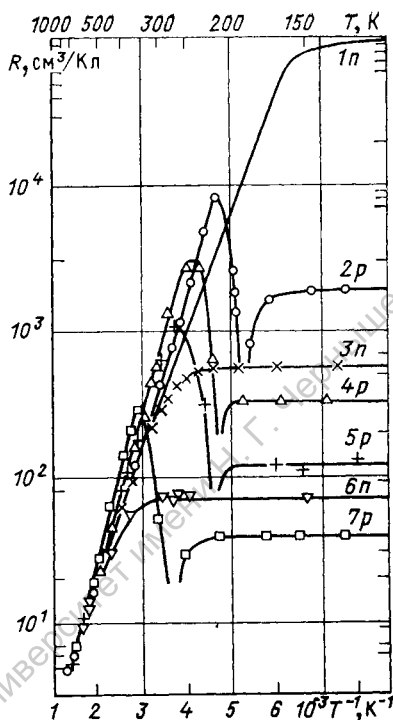


Рис. 6.8. Зависимость коэффициента $r = \frac{R}{1/(en)}$ от отношения ρ_I/ρ

Рис. 6.9. Температурная зависимость коэффициента Холла в антимониде индия для образцов n- и p-типа:

Концентрация доноров и акцепторов, см^{-3} ; $1n = 4,3 \times 10^{13}$; $2p = 4 \cdot 10^{16}$; $3n = 1,3 \cdot 10^{16}$; $4p = 2,2 \cdot 10^{16}$; $5p = 6 \cdot 10^{16}$; $6n = 1 \cdot 10^{17}$; $7p = 2 \cdot 10^{17}$



меси к полному сопротивлению ρ при рассеянии на ионах и на фоновых.

Для электронного полупроводника, у которого $n \gg p$, коэффициент Холла на основании (6.107) имеет вид:

$$R = -r \frac{1}{en}. \quad (6.112)$$

В то же время удельная проводимость такого полупроводника, определяемая дрейфовой подвижностью μ_d , равна $\sigma = en\mu_d$. Следовательно,

$$\frac{1}{r} |R| \sigma = \mu_d. \quad (6.113)$$

Но с помощью равенства

$$|R| \sigma = \mu_H \quad (6.114)$$

можно определить μ_H , имеющую размерность подвижности и вследствие этого называемую холловской подвижностью.

Из (6.113) и (6.114) следует, что

$$\mu_H = r\mu_d, \quad (6.115)$$

т. е. холловская подвижность, определяющая угол Холла, пропорциональна дрейфовой подвижности.

В случае, когда время релаксации $\tau = \text{const}$, т. е. не зависит от энергии,

$$r = 1, \quad (6.116)$$

поэтому

$$\mu_H = \mu_d. \quad (6.117)$$

Это имеет место у металлов и вырожденных полупроводников. У них коэффициент Холла R не зависит от механизма рассеяния.

Поскольку n и p являются функциями температуры, то R также зависит от температуры. В качестве примера на рис. 6.9 представлена температурная зависимость коэффициента Холла для образцов антимонида индия n - и p -типа. Как видно, в области низких температур концентрация электронов и дырок остается постоянной, что соответствует области истощения примеси. В области собственной электропроводности $\ln|R|$ линейно зависит от обратной температуры (кривые $1n$, $3n$ и $6n$). Так как у акцепторного полупроводника $R > 0$, то с ростом температуры R уменьшается, проходит через нуль и становится отрицательным. Температуру, при которой $R = 0$, называют температурой инверсии. Из кривых рис. 6.9 для образцов $2p$, $4p$, $5p$ и $7p$ следует, что температура инверсии, так же как и температура перехода к собственной электропроводности, тем выше, чем больше концентрация акцепторной примеси.

6.6. МАГНИТОРЕЗИСТИВНЫЙ ЭФФЕКТ

Искривление траектории носителей заряда в магнитном поле приводит к изменению составляющей тока, направленной вдоль вектора электрического поля, что эквивалентно изменению удельного сопротивления полупроводника. Это явление получило название эффекта магнитосопротивления (эффект Гаусса) или магниторезистивного эффекта.

Удельное сопротивление в отсутствие магнитного поля, когда направление плотности тока $\mathbf{J}$ совпадает с направлением напряженности электрического поля $\mathcal{E}$, определяется выражением вида

$$\rho = \frac{1}{\sigma} = \frac{\mathcal{E}}{J}. \quad (6.118)$$

При наличии магнитного поля, когда направление $\mathcal{E}$ не совпадает с направлением $\mathbf{J}$, в это уравнение вместо $\mathcal{E}$ следует подставить проекцию вектора $\mathcal{E}$ на вектор $\mathbf{J}$:

$$\rho_B = \frac{\mathcal{E}_J}{J} = \frac{(\mathcal{E} \mathbf{J})}{J^2}. \quad (6.119)$$

Выражение (6.103) для тока в слабых магнитных полях, полученное в линейном приближении относительно магнитной индукции $\mathbf{B}$, с учетом равенств (6.107), (6.47) и (6.51) можно переписать в виде

$$\mathbf{J} = \sigma \mathcal{E} - R \sigma^2 [\mathbf{B} \mathcal{E}]. \quad (6.120)$$

Оставаясь в рамках линейного приближения по магнитному полю и пренебрегая членами, содержащими B^2 , с учетом $\mathcal{E} [\mathbf{B}\mathcal{E}] = 0$ получаем:

$$J^2 = (\sigma \mathcal{E})^2. \quad (6.121)$$

Подстановка (6.120) и (6.121) в (6.119) дает:

$$\rho_B = \frac{\sigma \mathcal{E}^2}{(\sigma \mathcal{E})^2} = \frac{1}{\sigma} = \rho. \quad (6.122)$$

Таким образом, в слабом магнитном поле в линейном приближении сопротивление полупроводника не меняется.

Чтобы определить удельное сопротивление полупроводника в магнитном поле ρ_B , следует использовать квадратичное приближение, решая (6.25) для однородного полупроводника ($\nabla F = 0$), находящегося в слабом магнитном поле $e\tau_e B / m_n^* \ll 1$ в отсутствие градиента температуры ($\nabla T = 0$). С учетом этого будем иметь:

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \left\{ \frac{\mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B}\mathcal{E}] + \left(\frac{e\tau_e}{m_n^*} \right)^2 (\mathbf{B}\mathcal{E}) \mathbf{B}}{1 + \left(\frac{e\tau_e}{m_n^*} B \right)^2} \right\}. \quad (6.123)$$

Поскольку $\frac{e\tau_e}{m_n^*} B \ll 1$, то $\frac{1}{1 + \left(\frac{e\tau_e}{m_n^*} B \right)^2} \approx 1 - \left(\frac{e\tau_e}{m_n^*} B \right)^2$. В ре-

зультате получим:

$$\begin{aligned} \chi_n &= -\frac{e\hbar\tau_e}{m_n^*} \left[\mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B}\mathcal{E}] + \left(\frac{e\tau_e}{m_n^*} \right)^2 \{ (\mathbf{B}\mathcal{E}) \mathbf{B} - B^2 \mathcal{E} \} \right] = \\ &= -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B}\mathcal{E}] + \left(\frac{e\tau_e}{m_n^*} \right)^2 [\mathbf{B} [\mathbf{B}\mathcal{E}]] \right\}. \end{aligned} \quad (6.124)$$

Плотность тока для полупроводника с простой зонной структурой, содержащего один тип носителей заряда, например электроны, на основании выражения (6.32) запишется в виде

$$\mathbf{J}_n = \frac{e\hbar^2}{4\pi^3 m_n^{*2}} \int \frac{\partial f_0}{\partial E} \left\{ \tau_e (\mathbf{k}\mathcal{E}) + \left(\frac{e\tau_e}{m_n^*} \right)^2 \mathbf{k} [\mathbf{B}\mathcal{E}] + \frac{e^2 \tau_e^3}{m_n^{*2}} \mathbf{k} [\mathbf{B} [\mathbf{B}\mathcal{E}]] \right\} \mathbf{k} d\tau_{\mathbf{k}}. \quad (6.125)$$

Проводя интегрирование, находим с точностью до членов порядка B^2 :

$$\mathbf{J}_n = e n \mu_n \left\{ \mathcal{E} + r \mu_n [\mathbf{B}\mathcal{E}] + \frac{3}{2} r \mu_n^2 [\mathbf{B} [\mathbf{B}\mathcal{E}]] \right\}. \quad (6.126)$$

Для решения полученного уравнения разложим вектор $\mathcal{E}$ по некопланарным направлениям $\mathbf{J}$, $\mathbf{B}$ и $[\mathbf{B}\mathbf{J}]$:

$$\mathcal{E} = \alpha \mathbf{J} + \beta \mathbf{B} + \gamma [\mathbf{B}\mathbf{J}]. \quad (6.127)$$

Коэффициенты разложения нетрудно найти, подставив (6.126) в (6.127) и сравнив коэффициенты при векторах. В результате с точностью до B^2 получим:

$$\begin{aligned} \mathbf{E} = & \frac{1}{e n \mu_n} \left\{ \left[1 + \frac{3}{2} r \mu_n^2 \left(1 - \frac{\pi}{4} \right) B^2 \right] \mathbf{J} - \right. \\ & \left. - \frac{3}{2} r \left(1 - \frac{\pi}{4} \right) \mu_n^2 (\mathbf{J} \mathbf{B}) \mathbf{B} - r \mu_n [\mathbf{B} \mathbf{J}] \right\}, \end{aligned} \quad (6.128)$$

что при использовании формулы (6.119) приводит к

$$\frac{\rho_B - \rho}{\rho} = \frac{\Delta \rho}{\rho} = \frac{3}{2} r \left(1 - \frac{\pi}{4} \right) \mu_n^2 \left[B^2 - \frac{(\mathbf{J} \mathbf{B})^2}{J^2} \right] = A \mu_n^2 B^2 \sin^2 \alpha, \quad (6.129)$$

где коэффициент $A = \frac{3}{2} r \left(1 - \frac{\pi}{4} \right)$ зависит от механизма рассеяния носителей заряда; α — угол между направлениями векторов плотности тока $\mathbf{J}$ и индукции магнитного поля $\mathbf{B}$:

$$(\mathbf{J} \mathbf{B})^2 = J^2 B^2 \cos^2 \alpha. \quad (6.130)$$

Из (6.129) следует, что если $\mathbf{B}$ параллельно $\mathbf{J}$ ($\alpha = 0$), то сопротивление полупроводника со сферически симметричной зонной структурой не изменится, т. е. продольное магнитосопротивление равно нулю. Поперечное магнитосопротивление ($\mathbf{J} \perp \mathbf{B}$) в слабых полях пропорционально квадрату магнитной индукции и квадрату подвижности носителей заряда.

Чтобы найти магнитосопротивление при наличии носителей заряда обоих знаков, следует учесть дырочную составляющую плотности тока, которая по аналогии с уравнением (6.126) равна:

$$\mathbf{J}_p = e r \mu_p \left\{ \mathbf{E} - r \mu_p [\mathbf{B} \mathbf{E}] + \frac{3}{2} r \mu_p^2 [\mathbf{B} [\mathbf{B} \mathbf{E}]] \right\}. \quad (6.131)$$

Уравнение для полного тока тогда имеет вид:

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p = \sigma \left\{ \mathbf{E} + r \frac{n \mu_n^2 - p \mu_p^2}{n \mu_n + p \mu_p} [\mathbf{B} \mathbf{E}] + \frac{3}{2} r \frac{n \mu_n^3 + p \mu_p^3}{n \mu_n + p \mu_p} [\mathbf{B} [\mathbf{B} \mathbf{E}]] \right\} \quad (6.132)$$

и решается подобно (6.126).

В результате для полупроводника с простой зонной структурой, имеющего два типа носителей заряда, получим выражение

$$\frac{\Delta \rho}{\rho} = \frac{3}{2} r \frac{e}{\sigma} \left[(n \mu_n^3 + p \mu_p^3) - \frac{\pi}{4} \frac{(n \mu_n^2 - p \mu_p^2)^2}{n \mu_n + p \mu_p} \right] B^2 \sin^2 \alpha, \quad (6.133)$$

которое в случае $n = 0$ (или $p = 0$) сводится к (6.129) для полупроводника с одним типом носителей заряда.

В материалах, не обладающих сферической симметрией зон, изменение сопротивления в магнитном поле зависит от ориентации вектора $\mathbf{B}$ относительно кристаллографических осей. В общем случае наблюдается также эффект продольного магнитосопротивления.

Поворачивая образец в магнитном поле и измеряя $\Delta\rho/\rho$, можно получить важные сведения о параметрах зонной структуры данного материала.

Изменение сопротивления полупроводниковой пластины в магнитном поле существенно зависит от ее конфигурации (отношения длины к ширине). Возникающее в результате эффекта Холла поперечное электрическое поле препятствует изменению траектории носителей заряда под действием магнитного поля и снижает таким образом эффект магнитосопротивления. В частности, из формулы (6.133) следует, что отношение $\Delta\rho/\rho$ максимально в том случае, когда второе слагаемое обращается в нуль, что в соответствии

с (6.107) имеет место при $R = 0$, т. е. в отсутствие поля Холла. Учитывая это свойство, для достижения наибольшего изменения сопротивления в магнитном поле следует использовать короткие, широкие образцы с омическими контактами на длинных сторонах, шунтирующими ЭДС Холла. Оптимальной конструкцией является диск Корбино, у которого один контакт расположен в центре круговой пластины, а другой — по ее окружности. В таком образце ток протекает в радиальных направлениях, а ЭДС Холла полностью отсутствует. На рис. 6.10 проиллюстрировано изменение сопротивления в магнитном поле для прямоугольных пластин антимида индия различной конфигурации, а также для диска Корбино (R_0 , R_B — сопротивления образца при отсутствии и наличии магнитного поля).

Рассмотрим теперь случай сильного магнитного поля, когда $\frac{e\tau_e}{m_n^*} B \gg 1$. Разлагая векторную функцию χ_n (6.25) в ряд по параметру $m_n^*/e\tau_e B$, для однородного полупроводника и в отсутствие градиента температуры $\nabla_n f = \frac{E-F}{T} \nabla T - \nabla F = 0$ будем иметь:

$$\begin{aligned} \chi_n &= -\frac{e\hbar\tau_e}{m_n^*} \left\{ \frac{(\mathbf{B}\mathbf{E})\mathbf{B}}{B^2} + \frac{m_n^*}{e\tau_e B^2} [B\mathbf{E}] + \left(\mathbf{E} - \frac{(\mathbf{B}\mathbf{E})\mathbf{B}}{B^2} \right) \left(\frac{m_n^*}{e\tau_e B} \right)^2 \right\} = \\ &= -\frac{e\hbar\tau_e}{m_n^*} \left\{ (\mathbf{e}_B \mathbf{E}) \mathbf{e}_B + \frac{m_n^*}{e\tau_e B} [\mathbf{e}_B \mathbf{E}] + \left(\mathbf{E} - (\mathbf{e}_B \mathbf{E}) \mathbf{e}_B \right) \left(\frac{m_n^*}{e\tau_e B} \right)^2 \right\} = \\ &= -\frac{e\hbar\tau_e}{m_n^*} \mathbf{E}_B - \frac{\hbar}{B} [\mathbf{e}_B \mathbf{E}] - \frac{\hbar m_n^*}{e\tau_e B^2} \mathbf{E}_\perp, \end{aligned} \quad (6.134)$$

где $\mathbf{e}_B = \mathbf{B}/B$ — единичный вектор в направлении магнитного поля; $\mathbf{E}_B = (\mathbf{e}_B \mathbf{E}) \mathbf{e}_B$ — составляющая напряженности электрического поля в направлении магнитного поля; $\mathbf{E}_\perp = (\mathbf{E} - (\mathbf{e}_B \mathbf{E}) \mathbf{e}_B) = \mathbf{E} - \mathbf{E}_B$ —

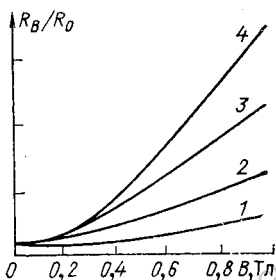


Рис. 6.10. Зависимость магнитосопротивления от отношения длины образца антимида индия к ширине: 1 — 2; 2 — 1; 3 — 0,5; 4 — 0 (диск Корбино)

$\mathbf{E}_\perp$ — составляющая напряженности электрического поля, перпендикулярная магнитному полю.

Решая уравнение для плотности тока, получаем:

$$\mathbf{J} = en\mu_n \mathbf{E}_B + \frac{en}{B} [\mathbf{e}_B \mathbf{E}] + \frac{4}{3r} \frac{en}{\mu_n B^2} \mathbf{E}_\perp. \quad (6.135)$$

Будем считать, что плотность тока $\mathbf{J}$ направлена по оси x , а напряженность магнитного поля — по оси z . Тогда

$$\left. \begin{aligned} J_x &= -\frac{en}{B} \mathcal{E}_y + \frac{4}{3r} \frac{en}{\mu_n B^2} \mathcal{E}_x = J; \\ J_y &= \frac{en}{B} \mathcal{E}_x + \frac{4}{3r} \frac{en}{\mu_n B^2} \mathcal{E}_y = 0. \end{aligned} \right\} \quad (6.136)$$

Решая эту систему уравнений с точностью до $1/B^2$ относительно $\mathcal{E}_x$ и $\mathcal{E}_y$, будем иметь:

$$\left. \begin{aligned} \mathcal{E}_x &= \frac{4}{3r} \frac{1}{en\mu_n} J; \\ \mathcal{E}_y &= -\frac{1}{en} JB. \end{aligned} \right\} \quad (6.137)$$

Отсюда найдем выражения для коэффициента Холла R и магнитосопротивления

$$R = -\frac{1}{en}; \quad (6.138)$$

$$\rho_B = \frac{(\mathcal{E}J)}{J^2} = \frac{4}{3r} \frac{1}{en\mu_n} = \frac{4}{3r} \rho. \quad (6.139)$$

В случае рассеяния на тепловых колебаниях решетки $r = 3\pi/8$, поэтому

$$\Delta\rho/\rho = 4/3r - 1 = 32/9\pi - 1 = 0,13. \quad (6.140)$$

Таким образом, в сильном магнитном поле коэффициент Холла не зависит от механизма рассеяния, а для магнитосопротивления наступает насыщение. При этом магнитосопротивление при насыщении зависит от механизма рассеяния.

В случае полупроводника со смешанной проводимостью плотность тока будет равна:

$$\mathbf{J} = e(n\mu_n + p\mu_p) \mathbf{E}_B + \frac{e}{B} (n - p) [\mathbf{e}_B \mathbf{E}] + \frac{4}{3r} \frac{e}{B^2} \left(\frac{n}{\mu_n} + \frac{p}{\mu_p} \right) \mathbf{E}_\perp. \quad (6.141)$$

На основании этого выражения, находя $J_x = J$ и $J_y = 0$, а затем определяя $\mathcal{E}_x$ и $\mathcal{E}_y$, получаем:

$$R = \frac{1}{e(p - n)}; \quad (6.142)$$

$$\rho_B = \frac{4}{3r} \frac{\frac{n}{\mu_n} + \frac{p}{\mu_p}}{e(p - n)}. \quad (6.143)$$

6.7. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ

К термоэлектрическим явлениям относятся три эффекта: Зеебека, Пельтье и Томсона. Рассмотрим эти эффекты.

1. Эффект Зеебека

Эффект Зеебека, или термоэлектрический эффект, заключается в том, что в замкнутой цепи, состоящей из двух различных полупроводников 1 и 2, места соединений которых находятся при различных температурах (рис. 6.11), возникает электрический ток, носящий название термоэлектрического. На концах такой разомкнутой цепи появляется разность потенциалов $\mathcal{E}$, которая носит название термоэлектродвижущей силы (термо-ЭДС).

Разность потенциалов, зависящая от разности температур и вида материала, характеризуется коэффициентом:

$$\alpha = \frac{d\mathcal{E}}{dT}. \quad (6.144)$$

Коэффициент α называется удельной термо-ЭДС и представляет собой термо-ЭДС, отнесенную к единичной разности температуры.

Если полупроводник нагрет неравномерно, то средняя энергия носителей заряда и их концентрация будут больше в той его области, где выше температура. Следовательно, градиент температуры в однородном полупроводнике приводит к градиенту средней энергии носителей заряда и градиенту их концентрации, вследствие чего возникает диффузионный поток носителей заряда и появляется электрический ток. В разомкнутой цепи в стационарном состоянии плотность тока в любой точке образца равна нулю. Это означает, что электрический ток, обусловленный градиентом температуры, компенсируется током, возникающим в электрическом поле при разделении зарядов. На образце возникает термоэлектродвижущая сила.

Вычисление термо-ЭДС проведем с помощью кинетического уравнения для однородного невырожденного полупроводника, имеющего сферические изоконцентрические поверхности для электронов $\left(E = \frac{\hbar^2 k^2}{2m_n^*}\right)$ и дырок $\left(E' = -\frac{\hbar^2 k'^2}{2m_p^*}\right)$. Если химический потенциал для электронов F , то для дырок он будет $F' = -E_g - F$, где E_g — ширина запрещенной зоны. Поскольку полупроводник невырожденный, то равновесная функция распределения для электронов $f_0 = e^{-(E-F)/kT}$ и дырок $f_{0p} = e^{(E'-F')/kT} = e^{(E'+E_g+F)/kT}$, а их концентрации соответственно равны: $n = N_c e^{-(E_c-F)/kT} = N_c e^{F/kT}$ и $p = N_v e^{-(F'+E_v)/kT} = N_v e^{-F/kT}$, так как $E_v = 0$,

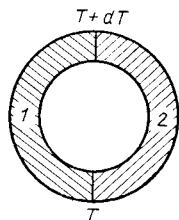


Рис. 6.11. Замкнутая цепь двух полупроводников

В этом случае с учетом того, что $\mathbf{B} = 0$ (нет магнитного поля), выражения для χ_n (6.25) и χ_p (6.26) будут иметь вид:

$$\chi_n = \frac{\hbar \tau_n}{m_n^*} \left[\frac{F - E}{T} \nabla T - \nabla (F - e\varphi) \right]; \quad (6.145)$$

$$\chi_p = - \frac{\hbar \tau_p}{m_p^*} \left[\frac{E' + E_g + F}{T} \nabla T - \nabla (F - e\varphi) \right]. \quad (6.146)$$

Вычисление плотности тока проведем для полупроводника в предположении, что рассеяние носителей заряда осуществляется на акустических колебаниях кристаллической решетки. В этом случае длина свободного пробега $l = v\tau$ не зависит от энергии носителей заряда (см. § 5.7) и время релаксации можно выразить через l , т. е.

$$\tau = \frac{l}{v} = \frac{m_n^* l}{\hbar} k^{-1}. \quad (6.147)$$

Электронную составляющую тока (6.32) с учетом (6.145), (6.147) и (6.40) можно представить в виде

$$\mathbf{J}_n = - \frac{e \hbar l}{3 \pi^2 m_n^*} \left[\left\{ \frac{F}{T} \nabla T - \nabla (F - e\varphi) \right\} \int_0^\infty \frac{\partial f_0}{\partial E} k^3 dk - \nabla T \int_0^\infty \frac{\partial f_0}{\partial E} E k^3 dk \right]. \quad (6.148)$$

Если ввести $E = \alpha kT$, то $k^3 dk = \frac{(2m_n^* kT)^3}{2\hbar^4} \alpha d\alpha$. При этом, учитывая (6.6) и (6.145), получаем:

$$\mathbf{J}_n = - \frac{4e n l}{3 (2\pi m_n^* kT)^{1/2}} \left\{ \frac{F}{T} \nabla T - \nabla (F - e\varphi) - 2k \nabla T \right\}. \quad (6.149)$$

Используя подвижность носителей заряда (6.56), выраженную через длину свободного пробега, (6.149) можно представить в виде

$$\mathbf{J}_n = n \mu_n \left\{ \nabla (F - e\varphi) + \left(2k - \frac{F}{T} \right) \nabla T \right\}. \quad (6.150)$$

Проведя расчет для дырочной составляющей тока, найдем:

$$\mathbf{J}_p = p \mu_p \left\{ \nabla (F - e\varphi) - \left(2k - \frac{F + E_g}{T} \right) \nabla T \right\}. \quad (6.151)$$

Следовательно, полная плотность тока

$$\begin{aligned} \mathbf{J} = n \mu_n \left\{ \nabla (F - e\varphi) + \left(2k - \frac{F}{T} \right) \nabla T \right\} + \\ + p \mu_p \left\{ \nabla (F - e\varphi) - \left(2k - \frac{F + E_g}{T} \right) \nabla T \right\}. \end{aligned} \quad (6.152)$$

Из соотношения (6.152) следует, что в полупроводнике при наличии градиента температуры возникают электронный и дырочный токи, которые обусловлены действием градиента электрохимического потенциала $\nabla (F - e\varphi)$ и градиента температуры ∇T .

Для нахождения термо-ЭДС необходимо определить разность потенциалов для разомкнутой цепи. Поскольку для разомкнутой

цепи $J = 0$, то напряженность электрического поля, обусловленного градиентом температуры и называемого термоэлектрическим, можно найти из (6.152), приравняв ток нулю. Но при измерении термо-ЭДС на границах полупроводника и измерительных металлических электродов существуют контактные разности потенциалов (см. § 9.4), которые не равны друг другу вследствие существующего температурного градиента. Измерительный прибор отметит ЭДС, равную термо-ЭДС полупроводника и разности контактных потенциалов измерительных электродов. Чтобы исключить последнее слагаемое, следует определить термо-ЭДС как градиент электрохимического потенциала $\nabla(F - e\varphi)$. Абсолютное значение $|\nabla F/e|$ будет равно разности контактных потенциалов в граничных точках, если считать, что полупроводник и металл на контакте находятся в термодинамическом равновесии. Поэтому удельная термо-ЭДС α определяется как

$$\alpha = \frac{|\nabla(\varphi - F/e)|}{|\nabla T|}. \quad (6.153)$$

Полагая ток в (6.152) равным нулю, получаем

$$\alpha = -\frac{k}{e} \frac{1}{n\mu_n + p\mu_p} \left\{ n\mu_n \left(2 - \frac{F}{kT} \right) - p\mu_p \left(2 - \frac{F + E_g}{kT} \right) \right\}. \quad (6.154)$$

Если воспользоваться формулами для концентрации электронов и дырок невырожденного полупроводника, то (6.154) можно записать в виде

$$\alpha = -\frac{k}{e} \left\{ \frac{\left[2 + \ln \frac{N_c}{n} \right] n\mu_n - \left[2 + \ln \frac{N_v}{p} \right] p\mu_p}{n\mu_n + p\mu_p} \right\}, \quad (6.155)$$

где N_c , N_v — эффективная плотность состояний в зоне проводимости и в валентной зоне.

Таким образом, термо-ЭДС полупроводника определяется двумя слагаемыми, каждое из которых соответствует вкладу, вносимому электронами и дырками, причем эти слагаемые имеют противоположные знаки.

Из выражения (6.155) следует, что термо-ЭДС отрицательна для электронного полупроводника. В этом случае на горячем торце образца возникает положительный объемный заряд, поскольку электроны диффундируют от горячего торца к холодному. В дырочном полупроводнике термо-ЭДС положительна, в нем, наоборот, на горячем торце возникает отрицательный объемный заряд. Таким образом, если полупроводник примесный, то направление внутреннего электрического поля и полярность термо-ЭДС определяются знаком основных носителей заряда и, следовательно, по знаку термо-ЭДС можно определить тип примесной проводимости исследуемого образца. В случае смешанной проводимости знак термо-ЭДС определяется не только соотношением концентраций носителей заряда, но и их подвижностью. В свою очередь, термо-ЭДС примесных полупроводников уменьшается с ростом содержания примеси. При нагревании примесного образца его термо-ЭДС также уменьшается.

Для собственного полупроводника $n = p = n_i$, $F = -E_g/2$ и (6.154) примет вид

$$\alpha = -\frac{k}{e} \frac{b-1}{b+1} (2 + E_g/2kT), \quad (6.156)$$

где $b = \mu_n/\mu_p$.

Таким образом, термо-ЭДС собственного полупроводника определяется лишь шириной запрещенной зоны и отношением подвижности электронов и дырок.

Изменение механизма рассеяния в полупроводнике, которое может иметь место при различных температурах, влияет на термо-ЭДС через подвижность и сказывается на первом слагаемом (в квадратных скобках) (6.155). Например, его значения для рассеяния на акустических колебаниях и ионах примеси в атомных кристаллах, оптических колебаниях в ионных кристаллах выше и ниже температуры Дебая соответственно равны 2; 4; 3 и 5/2.

На основании выражения (6.155) нельзя объяснить большое возрастание термо-ЭДС, которое обнаруживается иногда в чистом веществе при очень низких температурах. Это явление определяется электронно-фононным взаимодействием, которое получило название эффекта фононного увлечения.

При наличии градиента температуры в полупроводнике имеет место анизотропия в распространении фононов, так как градиент концентрации фононов вызывает их движение в направлении от горячего к холодному краю образца. При каждом акте электрон-фононного рассеяния, которое сопровождается поглощением фонона, электрон приобретает энергию и квазиимпульс фонона. Поскольку фононов с квазиимпульсом, направленным от горячего к холодному краю образца, больше, чем фононов с противоположно направленным квазиимпульсом, то электроны, поглощая фононный квазиимпульс, будут увлекаться фононным потоком. Вследствие этого электроны станут дрейфовать к холодной части образца. Дрейф электронов будет продолжаться до тех пор, пока в связи с их перераспределением не возникнет электрическое поле, которое уравновесит силу, действующую на электроны со стороны фононного потока. Эта разность потенциалов и представляет собой дополнительную термо-ЭДС, которая добавляется к термо-ЭДС, рассчитанной для случая отсутствия электронно-фононного взаимодействия. Эффект фононного увлечения приводит к значительному росту термо-ЭДС только при низких температурах, так как из-за взаимных столкновений распределение фононов медленно возвращается к равновесному. При более высоких температурах рассеяние фононов на фононах быстро восстанавливает равновесное распределение фононов и эффект увлечения исчезает.

Рассмотренные здесь энергетические соотношения для термо-ЭДС полупроводников хорошо согласуются с экспериментальными результатами, полученными для германия как при высоких, так и при низких температурах (рис. 6.12).

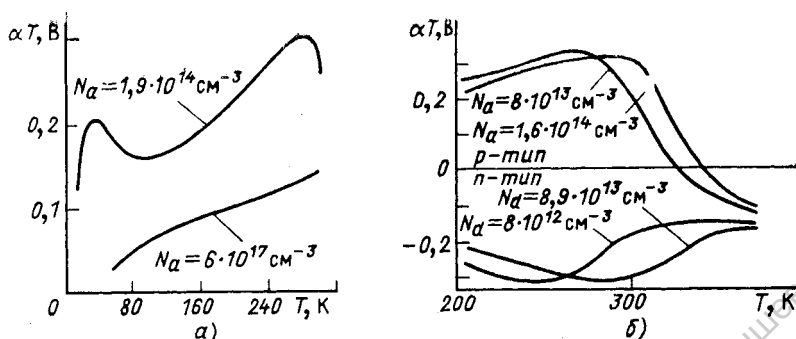


Рис. 6.12. Термо-ЭДС германия при низких (а) и высоких (б) температурах

У металлов концентрация электронов не зависит от температуры, а их средняя энергия:

$$\langle E \rangle = \frac{5}{3} F_0 \left[1 + \frac{5\pi^2}{12} \left(\frac{kT}{F_0} \right)^2 \right], \quad (6.157)$$

где F_0 — уровень Ферми при температуре абсолютного нуля.

Так как для металлов $F_0 \approx 5$ эВ, то второй член в (6.157) мал и энергия электронов в металле практически не зависит от температуры. В силу этого термо-ЭДС металлов мала по сравнению с полупроводниками. Для металлов $\alpha \approx 1 \div 10$ мкВ/К, а у полупроводников она достигает значения 100—1000 мкВ/К. Такое различие в термо-ЭДС приводит к тому, что при наличии в цепи полупроводника металла полная термо-ЭДС слабо зависит от свойств металла, а поэтому, как правило, для значений термо-ЭДС полупроводников не указывают, относительно какого металла она определена.

2. Эффект Пельтье

Эффект Пельтье заключается в том, что при прохождении электрического тока через контакт двух веществ в нем в зависимости от направления тока выделяется или поглощается теплота. Количество теплоты при эффекте Пельтье пропорционально плотности тока и времени:

$$Q_{\text{П}} = \Pi I t, \quad (6.158)$$

где Π — коэффициент Пельтье.

3. Эффект Томсона

Если вдоль однородного проводника существует градиент температуры, то при прохождении тока в объеме материала кроме тепла Джоуля выделяется (или поглощается) теплота, количество которой пропорционально количеству протекающего электричества и перепаду температур:

$$Q_T = It \int_T^{T'} \tau_T dT. \quad (6.159)$$

Коэффициент τ_T принято называть коэффициентом Томсона.

Коэффициенты α , Π и τ_T связаны между собой соотношением Томсона. Для их вывода проведем элементарное рассмотрение термоэлектрических эффектов для замкнутой цепи из двух проводников, как это изображено на рис. 6.11. Пусть между контактами полупроводников существует разность температур dT . Наличие градиента температуры вызывает термоэлектрическую силу, равную $d\mathcal{E} = \alpha dT$, которая обеспечит в данной цепи прохождение тока dI . При этом на одном из контактов в единицу времени выделяется теплота Пельтье в количестве $\Pi' dI$, а на другом поглощается ΠdI . В результате в цепи выделяется теплота Пельтье, и если Π' мало отличается от Π , то количество теплоты

$$dQ_{\Pi} = (\Pi' - \Pi) dI = \frac{d\Pi}{dT} dI dT. \quad (6.160)$$

Наличие электрического тока и градиента температуры вызовет появление теплоты за счет эффекта Томсона, равной за 1 с

$$dQ_T = (\tau_{T_2} - \tau_{T_1}) dI dT. \quad (6.161)$$

Эти термодинамические явления можно считать обратимыми, поскольку они зависят от направления тока и градиента температуры. Но в данной замкнутой цепи имеют место и необратимые процессы — теплопроводность и выделение теплоты Джоуля. В стационарных условиях теплопроводность не изменяет общего количества теплоты в системе, имеет место только перенос теплоты в неизменном количестве их одного участка в другой. Теплота Джоуля пропорциональна квадрату тока $(dI)^2$. Поэтому при малых значениях dI и dT эти необратимые эффекты можно не учитывать. На основании закона сохранения энергии имеем:

$$\alpha dI dT = \frac{d\Pi}{dT} dI dT + (\tau_{T_2} - \tau_{T_1}) dI dT, \quad (6.162)$$

откуда находим:

$$\alpha = \frac{d\Pi}{dT} + (\tau_{T_2} - \tau_{T_1}). \quad (6.163)$$

Если положить для горячего конца T' , Π' , τ_{T_2} , а для холодного T , Π , τ_{T_1} , то на основании второго закона термодинамики, по которому полное изменение энтропии dS замкнутой системы при обратимых процессах равно нулю, имеем:

$$\frac{\Pi' dI}{T'} - \frac{\Pi dI}{T} + dI \int_T^{T'} \frac{\tau_{T_2} - \tau_{T_1}}{T} dT = 0. \quad (6.164)$$

Если за начало отсчета энтропии взять ее значение на горячем конце ($\Pi' dI/T' = 0$), то, продифференцировав выражение (6.164) по T , получим:

$$\frac{d}{dT} \left(\frac{\Pi}{T} \right) - \frac{\tau_{T_2} - \tau_{T_1}}{T} = 0, \quad (6.165)$$

откуда найдем:

$$\frac{d\Pi}{dT} - (\tau_{T_2} - \tau_{T_1}) = \frac{\Pi}{T}. \quad (6.166)$$

Сравнивая (6.163) и (6.166), видим, что

$$\Pi = \alpha T. \quad (6.167)$$

Если продифференцируем равенство (6.167) по T , найдем:

$$\frac{d\alpha}{dT} = \frac{1}{T} \frac{d\Pi}{dT} - \frac{\Pi}{T^2} = \frac{1}{T} \left(\frac{d\Pi}{dT} - \frac{\Pi}{T} \right). \quad (6.168)$$

Учитывая выражение (6.166), получаем:

$$\tau_{T_2} - \tau_{T_1} = \tau_T = T \frac{d\alpha}{dT}. \quad (6.169)$$

Таким образом, коэффициенты α , Π , τ_T , характеризующие термоэлектрические свойства данного вещества, связаны между собой термодинамическими соотношениями. Обычно проводится экспериментальное измерение α , а трудно измеряемые коэффициенты Π и τ_T определяются на основе этих соотношений.

6.8. ТЕПЛОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ

Как известно, все тела проводят теплоту. Это явление носит название *теплопроводности*. В том случае, если в веществе нет тепловых процессов, приводящих к появлению тока, т. е. $\mathbf{J} = 0$, то имеющийся в веществе градиент температуры ∇T вызовет поток энергии $\mathbf{W}$ в направлении, противоположном ∇T :

$$\mathbf{W} = -\kappa \nabla T. \quad (6.170)$$

Здесь $\kappa = \frac{|\mathbf{W}|}{|\nabla T|}$ есть коэффициент теплопроводности, который численно равен количеству энергии, проходящей в единицу времени через единичное сечение образца, на концах которого создана разность температур в один градус.

Перенос тепловой энергии в твердом теле осуществляется свободными носителями заряда и фононами. Теплопроводность, обусловленную движением носителей заряда, называют *электронной* (или *дырочной*) *теплопроводностью*. Она характеризуется коэффициентом теплопроводности κ_e , который в сильной степени зависит от свойств полупроводника и температуры. Теплопроводность, обусловленную колебаниями решетки, называют *решеточной*, или *фононной*, *теплопроводностью*. Она обозначается κ_L . Таким образом, полная теплопроводность твердого тела складывается из решеточной κ_L и электронной (или дырочной) κ_e теплопроводности, т. е.

$$\kappa = \kappa_L + \kappa_e. \quad (6.171)$$

Теплопроводность диэлектриков в основном является фононной; у них $\kappa_L \gg \kappa_e$. В металлах высокой чистоты вклад решеточной теплопроводности обычно мал, поэтому $\kappa_e \gg \kappa_L$. В полупроводниках

нередко решеточная теплопроводность такого же порядка, что и электронная (или дырочная).

Решеточная теплопроводность определяется потоком фононов, движущихся от горячего конца образца к холодному. Если для фононов ввести длину свободного пробега l_L , то коэффициент фононной теплопроводности может быть определен на основе кинетической теории. При этом выражение для κ_L имеет вид:

$$\kappa_L = C v_{\text{зв}} l_L / 3, \quad (6.172)$$

где C — теплоемкость решетки; $v_{\text{зв}}$ — фононная скорость (скорость звука).

В линейном образце длиной $\mathcal{L}$, перепад температуры между концами которого равен $T_1 - T_2$, в единицу времени через единичную площадку проходит теплота в количестве

$$Q = \frac{\kappa_L}{\mathcal{L}} (T_1 - T_2). \quad (6.173)$$

Воспользовавшись для κ_L (6.172), получим:

$$Q = C (T_1 - T_2) l_L v_{\text{зв}} / 3 \mathcal{L}. \quad (6.174)$$

Определение теплопроводности κ_e , обусловленной движением носителей заряда, проведем для невырожденного донорного полупроводника, в котором рассеяние носителей заряда осуществляется на акустических колебаниях. Для этого вычислим плотность потока энергии, переносимого электронами:

$$\begin{aligned} \mathbf{W} &= \frac{1}{4\pi^3} \int_{(V_B)} f_1^{(n)}(\mathbf{k}) (E - e\varphi) \mathbf{v}_n d\tau_{\mathbf{k}} = \\ &= - \frac{e\hbar}{4\pi^3 m_n^*} \int_{(V_B)} \frac{\partial f_0}{\partial E} (\chi_n \mathbf{k}) (E - e\varphi) \mathbf{k} d\tau_{\mathbf{k}} \end{aligned} \quad (6.175)$$

в предположении, что

$$\mathbf{J}_n = 0. \quad (6.176)$$

Решая уравнение (6.175) с учетом (6.25) для случая $\mathbf{B} = 0$ и $\mathbf{J}_n = 0$, получаем

$$\mathbf{W} = -n\mu_n \frac{2k^2 T}{e} \nabla T. \quad (6.177)$$

В этом выражении множитель, стоящий перед ∇T , и есть коэффициент теплопроводности κ_e , равный:

$$\kappa_e = 2k^2 T n \mu_n / e = 2 (k/e)^2 \sigma T. \quad (6.178)$$

Следовательно, *теплопроводность вещества, обусловленная свободными носителями заряда, определяется температурой и его удельной проводимостью.*

Из (6.178) найдем отношение

$$\kappa_e/\sigma = 2 (k/e)^2 T, \quad (6.179)$$

которое выражает закон Видемана—Франца.

Проведя подобные вычисления теплопроводности для невырожденного полупроводника с двумя типами носителя заряда, получим

$$\begin{aligned} \kappa_e = & \frac{2k^2T}{e} n\mu_n + \frac{2k^2T}{e} p\mu_p + \\ & + \frac{n\mu_n p\mu_p (E_g + 4kT)^2}{eT (n\mu_n + p\mu_p)}. \end{aligned} \quad (6.180)$$

Установлено, что теплопроводность твердых тел, как следует из рис. 6.13, сначала растет с ростом температуры, а затем убывает при высоких температурах.

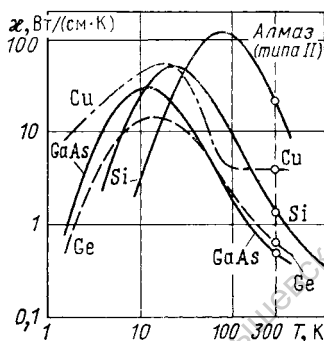


Рис. 6.13. Температурная зависимость теплопроводности германия, кремния, арсенида галлия, меди и алмаза (типа II)

6.9. ЭЛЕКТРОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ В СИЛЬНОМ ЭЛЕКТРИЧЕСКОМ ПОЛЕ

При выводе кинетического уравнения предполагалось, что время релаксации τ не зависит от электрического поля. Затем при помощи кинетического уравнения для невырожденного полупроводника со сферическими изонергетическими поверхностями было показано, что плотность тока подчиняется закону Ома, т. е. пропорциональна напряженности электрического поля:

$$\mathbf{J} = e n \mu_0 \mathbf{E} = \sigma_0 \mathbf{E}, \quad (6.181)$$

где μ_0 , σ_0 — подвижность и удельная проводимость при слабом электрическом поле.

При столкновении с акустическим фононом электрон может либо передать энергию фонону, либо получить энергию от фонона. Пусть при этом изменение энергии электрона будет ΔE .

Если из общего числа столкновений подсчитать число столкновений, идущих с передачей энергии фонону, то согласно (5.99) на одно такое столкновение приходится $2N_q + 1$ столкновение. Поэтому относительная потеря энергии электрона при столкновении составляет

$$\delta = \frac{\Delta E}{kT} = \frac{\hbar\omega_q}{kT} \frac{1}{2N_q + 1} = \frac{\hbar\omega_q}{kT} \operatorname{th} \frac{\hbar\omega_q}{2kT}. \quad (6.182)$$

Следовательно, изменение энергии электрона при столкновении с фононом равно $\Delta E = kT\delta$. Соударение будет упругим в том случае, если энергия, набираемая электроном в электрическом поле

на длине свободного пробега $e\mathcal{E}l$, будет значительно меньше энергии ΔE , теряемой в среднем при одном столкновении с решеткой

$$\frac{e\mathcal{E}l}{\Delta E} \ll 1.$$

Это означает, что закон Ома выполняется в том случае, когда

$$e\mathcal{E}l \ll kT\delta, \quad (6.183)$$

или

$$\mathcal{E} \ll \delta \frac{kT}{el}. \quad (6.184)$$

При этом условии распределение скоростей носителей заряда остается почти максвелловским с температурой, равной температуре решетки. В этом случае на хаотическое движение носителей заряда накладывается медленный их дрейф, обусловленный действием слабого внешнего электрического поля. При этом скорость дрейфа носителей заряда v значительно меньше их тепловой скорости v_0 , т. е.

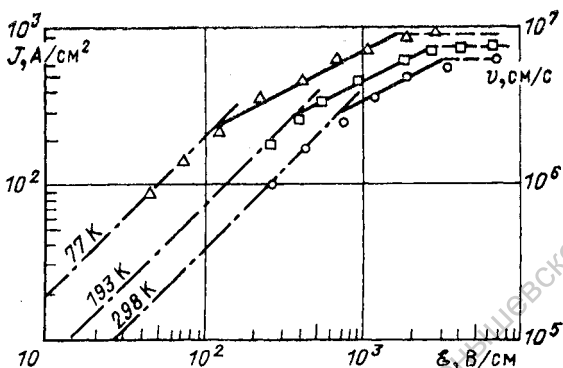
$$v \ll v_0. \quad (6.185)$$

В сильных электрических полях имеет место нарушение закона Ома. Это может происходить по двум причинам. Во-первых, в сильных электрических полях скорость дрейфа носителей заряда становится соизмеримой с их тепловой скоростью. Это скажется на процессе рассеяния, что проявляется в изменении времени релаксации τ , а следовательно, и подвижности μ . Сюда необходимо отнести явление «разогрева» электронно-дырочного газа и эффект Ганна. Во-вторых, сильные электрические поля могут вызвать изменение концентрации носителей заряда. При этом рост концентрации носителей заряда может обуславливаться рядом причин. Такими эффектами являются ударная ионизация, эффект Зинера и электростатическая ионизация.

Рассмотрим электропроводность полупроводника в сильном электрическом поле. В сильных электрических полях, для которых не выполняется условие (6.184), носители заряда при столкновении с акустическими фононами не успевают отдавать кристаллической решетке набираемую ими в поле энергию. В результате средняя энергия носителей заряда повышается и электронно-дырочный газ «разогревается». Но с ростом энергии носителей заряда увеличивается их скорость, а значит, уменьшается время свободного пробега и, следовательно, увеличивается вероятность рассеяния носителей заряда на тепловых колебаниях решетки. В силу этого рост энергии носителей заряда с возрастанием напряженности поля будет продолжаться до тех пор, пока электронно-фононное взаимодействие не усилится настолько, что установится стационарное состояние, при котором энергия, приобретаемая носителем заряда в электрическом поле за единицу времени $e\mathcal{E}v$, сравнивается с энергией, теряемой этим

Рис. 6.14. Зависимость плотности тока и дрейфовой скорости от напряженности внешнего электрического поля в электронном германии:

--- экстраполяция на область слабых полей;
— закон $\mathcal{E}^{1/2}$;
--- закон $\mathcal{E}^0$



носителем заряда за одну секунду в результате столкновения $\Delta E/\tau = kT\delta/\tau$, т. е.

$$e\mathcal{E}v = \delta kT/\tau. \quad (6.186)$$

При рассеянии на тепловых колебаниях решетки согласно (5.93) длина свободного пробега l не зависит от энергии, поэтому время свободного пробега $\tau = l/v_0$. Но дрейфовая скорость носителей заряда на основании (6.186) равна:

$$v = \delta kT/e\mathcal{E}\tau = \delta kTv_0/e\mathcal{E}l. \quad (6.187)$$

В то же время дрейфовая скорость

$$v = \frac{e\tau}{m^*} \mathcal{E} = \frac{el}{m^*v_0} \mathcal{E}. \quad (6.188)$$

Из сравнения (6.187) и (6.188) следует, что

$$(e\mathcal{E}l)^2 = 2\delta kTm^*v_0^2/2. \quad (6.189)$$

Но так как $m^*v_0^2/2 = kT$, то из (6.189) найдем критическую напряженность поля

$$\mathcal{E}_{кр} = \frac{kT}{el} (2\delta)^{1/2}, \quad (6.190)$$

при которой начинается разогрев электронно-дырочного газа. При этом, как следует из (6.190), напряженность критического поля уменьшается при понижении температуры. Это согласуется с экспериментальными данными, которые представлены на рис. 6.14.

Если воспользоваться равенствами (6.187) и (6.190), получим:

$$v = (\delta/2)^{1/2} v_0. \quad (6.191)$$

Следовательно, разогрев электронно-дырочного газа будет происходить в том случае, когда скорость, приобретенная носителем заряда во внешнем электрическом поле, станет сравнимой по величине с его тепловой скоростью.

Из (6.189) можно определить энергию носителей заряда, выраженную через тепловую скорость

$$\left(\frac{m^* v_0^2}{2}\right)^2 = \frac{(e\mathcal{E}l)^2}{2\delta}, \quad (6.192)$$

откуда

$$v_0 = \left[\frac{2e\mathcal{E}l}{m^* (2\delta)^{1/2}} \right]^{1/2}. \quad (6.193)$$

Подставляя (6.193) в (6.191), получаем:

$$v = \left(\frac{\delta}{2}\right)^{1/2} \left[\frac{2e\mathcal{E}l}{m^* (2\delta)^{1/2}} \right]^{1/2} = \left[\frac{el}{m^*} \left(\frac{\delta}{2}\right)^{1/2} \right]^{1/2} \mathcal{E}^{1/2}. \quad (6.194)$$

Таким образом, в сильных электрических полях, когда напряженность электрического поля больше критического, т.е.

$$\mathcal{E} > \mathcal{E}_{кр} \text{ или } \mathcal{E} > \frac{kT}{el} (2\delta)^{1/2}, \quad (6.195)$$

дрейфовая скорость движения носителей заряда будет увеличиваться как

$$v \sim \mathcal{E}^{1/2}, \quad (6.196)$$

а время релаксации соответственно будет уменьшаться:

$$\tau = \frac{l}{v_0 + v} \sim \mathcal{E}^{-1/2}. \quad (6.197)$$

Отсюда следует, что

$$\mu_n = \frac{e\tau}{m^*} = \mu_0 \mathcal{E}^{-1/2}, \quad (6.198)$$

а

$$\sigma = en\mu_n = \sigma_0 \mathcal{E}^{-1/2} \text{ и } J = \sigma \mathcal{E} = \sigma_0 \mathcal{E}^{1/2}, \quad (6.199)$$

Следовательно, при напряженности внешнего электрического поля $\mathcal{E} > \mathcal{E}_{кр}$ закон Ома не выполняется. Например, в германии отклонения от омической проводимости при комнатной температуре наступают при напряженности поля около 700 В/см (рис. 6.14).

В области еще более сильных электрических полей энергия носителей заряда увеличивается настолько, что возможно рассеяние их на оптических колебаниях. В этом случае при столкновении с оптическим фоном энергия носителя изменяется на значение, равное $\hbar\omega_0$, и скорость потерь энергии сильно возрастает. В результате устанавливается новое стационарное состояние, при котором энергия $e\mathcal{E}v$, набираемая носителем заряда в поле на единицу времени, равна энергии $\Delta E/\tau$, теряемой этим носителем заряда за то же время, т.е.

$$e\mathcal{E}v = \frac{\Delta E}{\tau}. \quad (6.200)$$

Но энергия, которую носитель заряда теряет при одном столкновении, согласно (6.182)

$$\Delta E = \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2kT}, \quad (6.201)$$

поэтому

$$e\mathcal{E}v = \frac{v_0}{l} \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2kT}. \quad (6.202)$$

Если из (6.188) найти напряженность электрического поля $\mathcal{E}$ и подставить ее в (6.202), получим дрейфовую скорость носителей заряда

$$v = \left(\frac{\hbar\omega_0}{m^*} \operatorname{th} \frac{\hbar\omega_0}{2kT} \right)^{1/2}. \quad (6.203)$$

Предельное значение скорости дрейфа

$$v_s = (\hbar\omega_0/m^*)^{1/2}. \quad (6.204)$$

Таким образом, в пределе плотность тока, текущего через образец, стремится к насыщению и равна

$$J_s = env_s = en(\hbar\omega_0/m^*)^{1/2}. \quad (6.205)$$

Это значит, что

$$\sigma = \frac{J_s}{\mathcal{E}} = \sigma_0 \mathcal{E}^{-1}. \quad (6.206)$$

Из (6.202) и (6.203) следует, что при увеличении напряженности электрического поля до

$$\mathcal{E}'_{кр} = \frac{v_0}{el} \left(m^* \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2kT} \right)^{1/2} \quad (6.207)$$

в ковалентных кристаллах наступает насыщение тока, а следовательно, и насыщение дрейфовой скорости.

При $\mathcal{E} > \mathcal{E}'_{кр}$ ток и дрейфовая скорость не зависят от напряженности внешнего электрического поля. Оценки показывают, что насыщение тока и дрейфовой скорости наступает в полях порядка 10^4 В/см, а дрейфовая скорость достигает $v \approx 10^7$ см/с. Это согласуется с экспериментальными данными рис. 6.15, на котором представлена зависимость скорости электронов и дырок от электрического поля для чистых германия и кремния. Из данных этого рисунка следует, что при слабых полях выполняется закон Ома. Начиная с критического значения напряженности поля, зависящего от температуры, ток пропорционален $\mathcal{E}^{1/2}$. При очень высоких полях ток не зависит от поля.

В полярных кристаллах рассеяние на оптических колебаниях становится преобладающим уже при низких значениях энергии носителей заряда. Во внешних электрических полях, при которых

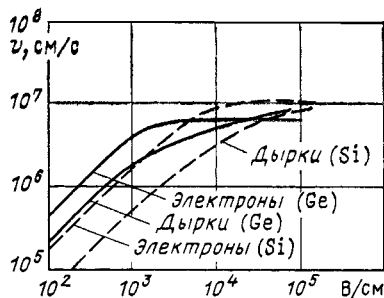


Рис. 6.15. Зависимость скорости носителей заряда от напряженности электрического поля для чистых германия и кремния ($T = 300$ К)

энергия электрона $E \ll \hbar \omega_0$, подвижность, как и в ковалентных полупроводниках, уменьшается с ростом напряженности поля. В очень сильных полях, когда энергия электрона $E \gg \hbar \omega_0$, вместо уменьшения подвижности и насыщения дрейфовой скорости имеет место быстрый их рост и энергия носителей заряда увеличивается. Это явление получило название «убегания» носителей заряда.

6.10. ЭФФЕКТ ГАННА

В полупроводниках, зона проводимости которых имеет более одного минимума энергии, электрон с волновым вектором k , соответствующим одному из минимумов, при рассеянии может оказаться в состоянии с волновым вектором k' , принадлежащим другому минимуму. В результате такого рассеяния будет иметь место переброс электронов из одного минимума в другой минимум зоны проводимости. Такой вид рассеяния получил название междолинного.

Междолинное рассеяние носителей заряда в определенных условиях может приводить к возникновению колебаний тока с частотой порядка 10^{10} Гц при приложении к однородному полупроводнику сильного постоянного электрического поля. Это явление, названное эффектом Ганна, впервые наблюдалось на GaAs.

На рис. 6.16 изображена энергетическая структура арсенида галлия в направлении оси $[100]$.

Существенным здесь является наличие минимумов A и B , разделенных зазором $\Delta E = 0,36$ эВ, в которых эффективные массы электронов различны. В области низшей долины A электроны легкие, с эффективной массой $m_1^* = 0,068 m_0$, имеют высокую подвижность [$\mu_1 \approx 4000 + 8000$ см²/(В·с)]. В области высокой долины B электроны тяжелые с $m_2^* = 1,2 m_0$ и обладают низкой подвижностью [$\mu_2 \approx 100$ см²/(В·с)]. Плотность состояний в верхней долине примерно в 70 раз выше, чем в нижней.

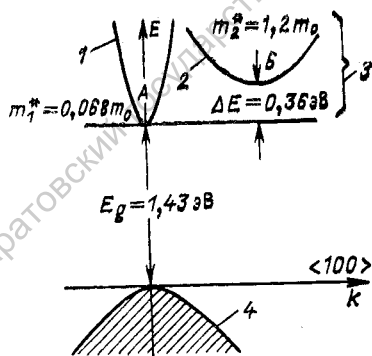


Рис. 6.16. Схематическое изображение структуры зон арсенида галлия n -типа в направлении $[100]$:

1 — нижний минимум (A); 2 — верхний минимум (B); 3 — зона проводимости; 4 — валентная зона

При достаточно низких температурах и в слабом внешнем электрическом поле энергия электронов E значительно меньше энергетического зазора $kT \ll \Delta E$, т. е. $E \ll \Delta E$, поэтому электроны в основном будут занимать энергетические уровни в низшей долине зоны проводимости (рис. 6.17, а), где они имеют высокую подвижность, малую эффективную массу и малую плотность состояний. В этом случае плотность тока определяется концентрацией элект-

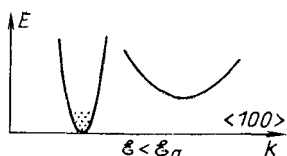


Рис. 6.17. Распределение электронов при различных значениях напряженности электрического поля

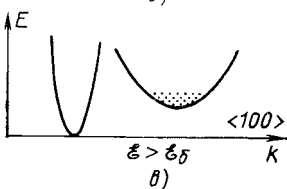
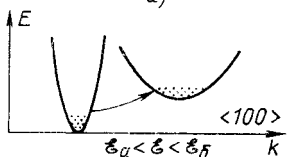
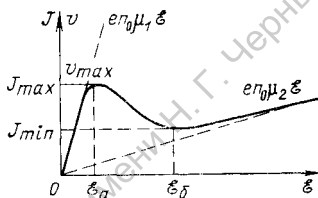


Рис. 6.18. Зависимость плотности тока от напряженности электрического поля для двухдолинного полупроводника



тронов в нижней долине $n_1 = n_0$ (в верхней долине $n_2 = 0$), их подвижностью μ_1 и будет линейно возрастать с ростом напряженности поля до некоторого критического значения $\mathcal{E}_a$ (рис. 6.18):

$$J = e n_0 \mu_1 \mathcal{E}. \quad (6.208)$$

По мере возрастания напряженности электрического поля средняя энергия электронов повышается и при $E > \Delta E$ становится возможным переход электронов в верхнюю долину Б (рис. 6.17, б), где они будут иметь низкую подвижность, большую эффективную массу и высокую плотность состояний (широкий максимум), а суммарная концентрация электронов будет $n_0 = n_1 + n_2$. Таким образом, с ростом напряженности поля (от $\mathcal{E}_a$ до $\mathcal{E}_b$ — рис. 6.18) будет иметь место уменьшение подвижности электронов, а следовательно, уменьшение плотности тока и на вольт-амперной характеристике появится падающий участок.

При $\mathcal{E} > \mathcal{E}_b$ все электроны будут находиться в верхней долине $n_2 = n_0$, а в нижней долине $n_1 = 0$ (см. рис. 6.17, в) и плотность тока:

$$J = e n_0 \mu_2 \mathcal{E}. \quad (6.209)$$

Рассмотрим подробнее механизм такой неустойчивости. Предположим, что к образцу длиной $\mathcal{L}$ приложено внешнее напряжение. В однородном полупроводнике электрическое поле примерно одинаково по всей длине образца. Но если в образце имеется локальная неоднородность с повышенным сопротивлением, то напряженность электрического поля в этом месте образца будет несколько выше. Следовательно, критическое значение $\mathcal{E}_{кр} = \mathcal{E}_a$ при повышении напряженности поля возникнет в первую очередь в этом сечении

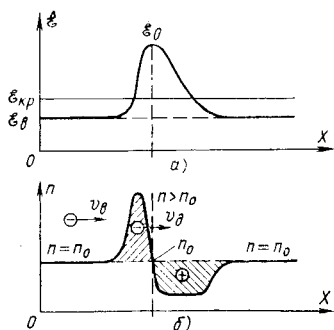


Рис. 6.19. Распределение электрического поля (а) и плотности электронов (б) вдоль образца в случае стабильного домена сильного поля, движущегося со скоростью U_d

неоднородным и образуется область сильного электрического поля, которая носит название электрического домена (рис. 6.19, а).

Область тяжелых электронов под действием электрического поля домена будет перемещаться вдоль образца с относительно низкой скоростью, обусловленной низкой подвижностью тяжелых электронов. Справа и слева от зоны тяжелых электронов будут двигаться легкие электроны с большой скоростью. Слева они будут догонять эту зону, и в результате образуется область повышенной концентрации электронов — область отрицательного объемного заряда. Справа от этой зоны легкие электроны будут уходить вперед, поэтому образуется область, обедненная электронами, — область положительного объемного заряда. Следовательно, в пределах области сильного электрического поля домена на кривой распределения концентрации электронов имеется обедненный участок с $n < n_0$, соответствующий фронту домена, и обогащенный участок с $n > n_0$, соответствующий спаду домена (рис. 6.19, б).

Поскольку внутри домена напряженность поля $\mathcal{E}_d$ сильно возрастает, возрастает в нем и скорость движения электронов v_d . Вне домена напряженность поля $\mathcal{E}_в$ резко уменьшается, поэтому скорость движения электронов падает. Нарастание напряженности поля в домене прекратится тогда, когда скорость движения домена v_d станет равной дрейфовой скорости электронов вне домена $v_в$, т. е.

$$v_d = v_в \quad (6.210)$$

или

$$\mu_2 \mathcal{E}_d = \mu_1 \mathcal{E}_в. \quad (6.211)$$

При этом установившаяся скорость движения домена v_d будет меньше максимальной скорости движения электронов, которую они имеют при $\mathcal{E}_а$. Поэтому в момент подключения к образцу напряжения

образца. Как только напряженность поля в области локальной неоднородности достигнет критического значения $\mathcal{E}_а$, электроны будут переходить в верхнюю долину Б и в этой области образца с повышенной напряженностью поля появятся тяжелые электроны. Подвижность электронов в этой части образца уменьшается и сопротивление ее возрастает. Это приводит к возрастанию напряженности поля в этом месте образца, что в свою очередь вызывает более интенсивный переход электронов в верхнюю долину. Но так как напряжение, приложенное к образцу, не меняется, то напряженность поля справа и слева от этой области образца будет падать. В результате распределение электрического поля станет резко

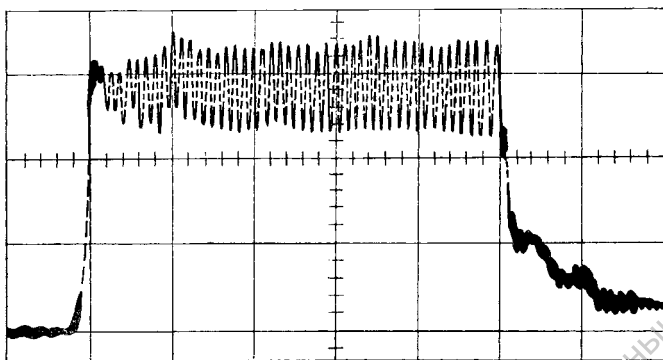


Рис. 6.20. Форма импульса тока при подаче на образец арсенида галлия длиной $2,5 \cdot 10^{-3}$ см импульса напряжения амплитудой 16 В и длительностью 16 нс. Частота переменной составляющей 4,5 ГГц

(через $t = t_1$) ток будет иметь максимальное значение I_{max} , определяемое v_{max} . Сразу же начнется процесс образования домена, и поскольку этот процесс кратковременный, так как постоянная времени, связанная с междолинным переходом электрона, порядка 10^{-12} с, ток очень быстро падает до значения I_{min} , определяемого скоростью установившегося движения домена v_d :

$$I_{min} = se n_0 v_d, \quad (6.212)$$

где s — площадь сечения образца.

Минимальное значение тока сохраняется в течение всего времени движения домена вдоль образца. Как установлено, это время определяется длиной образца и скоростью движения домена:

$$T = \frac{L}{v_d}. \quad (6.213)$$

Домен движется от катода к аноду. По достижении анода область сильного поля выходит из образца и ток в образце начинает возрастать. Как только ток в образце достигнет значения I_{max} , происходит образование нового домена и ток падает до I_{min} . В результате движения домена по кристаллу во внешней цепи появляется импульс тока. Скважность импульсов тока определяется временем T прохождения домена. При длине образца 50 мкм частота колебаний тока должна составлять около 2 ГГц. Примерный вид этих колебаний показан на рис. 6.20.

Поскольку эффект образования домена представляет собой объемную неустойчивость, он не зависит от внешних элементов схемы. Это значит, что скорость движения домена не зависит от величины внешнего напряжения и домен передвигается со скоростью дрейфа порядка 10^7 см/с.

Несмотря на то что в кристалле может быть несколько неоднородностей, на которых могут формироваться домены, в кристалле

существует только один домен. Другими словами, возникновение домена происходит только на одной из неоднородностей. После исчезновения домена новый домен может возникнуть на другой неоднородности. Разное расположение неоднородностей в кристалле определяет разные пролетные времена доменов, т. е. разный период колебаний. Поэтому для наблюдения эффекта Ганна необходимы чистые и очень однородные образцы. В пластиночных образцах домены зарождаются в области повышенного поля вблизи катода, обусловленной неоднородностью уровня легирования, которая возникает в процессе изготовления электродов. Важно также, чтобы расстояние между минимумами A и B зоны проводимости не было весьма велико, ибо для перехода электрона во вторую долину потребуются поле большой напряженности, при котором возможно увеличение концентрации электронов за счет ударной ионизации или туннельного эффекта.

6.11. УДАРНАЯ ИОНИЗАЦИЯ

В сильном электрическом поле, если оно возрастает примерно до 10^5 В/см, электроны (или дырки) приобретают энергию, достаточную для ионизации атомов вещества. В результате этого возникают электронно-дырочные пары (рис. 6.21), которые в свою очередь ускоряются полем и генерируют дополнительные носители заряда. Этот процесс создания носителей заряда носит название ударной ионизации. Наблюдать, однако, это явление возможно лишь в p - n переходах, так как создать поля 10^5 — 10^6 В/см в однородном полупроводнике чрезвычайно трудно при разумных толщинах последнего. Исключением является ударная ионизация примесных атомов, которая требует небольших полей и наблюдается при низких температурах, когда атомы примеси не ионизованы. Ударная ионизация доноров и акцепторов — элементов пятой и третьей групп — в германии происходит при полях, приблизительно равных 5—10 В/см. Например, из рис. 6.22, на котором представлена

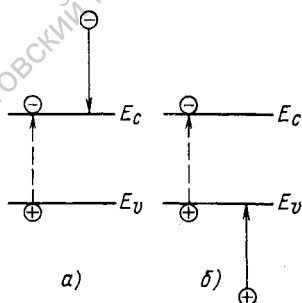


Рис. 6.21. Два возможных варианта ударной ионизации

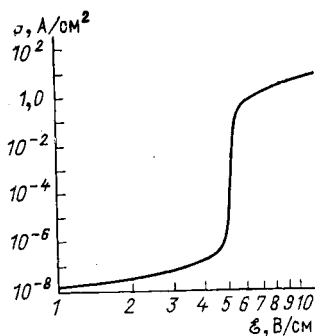


Рис. 6.22. Ударная ионизация в электронном германии при $T = 4,2$ К

зависимость плотности тока от электрического поля в электронном германии при 4,2 К, следует, что при $\mathcal{E} \approx 5$ В/см происходит резкое скачкообразное возрастание плотности тока, вызванное ударной ионизацией атомов примеси.

Скорость генерации электронно-дырочных пар G вследствие ударной ионизации дается в виде

$$G = \alpha_n \mu_n + \alpha_p \mu_p, \quad (6.214)$$

где α_n и α_p — коэффициент ударной ионизации электронов и дырок соответственно. Он определяется как число электронно-дырочных пар, генерируемых электроном (дыркой) на единице пути перемещения. Коэффициент ударной ионизации как электронов α_n , так и дырок α_p сильно зависит от значения напряженности электрического поля. На рис. 6.23 приведены экспериментально измеренные значения коэффициентов ударной ионизации для германия, кремния, арсенида галлия и фосфида галлия.

Коэффициент ударной ионизации $\alpha(E)$ определяется соотношением

$$\alpha(E) = \frac{1}{v} \int_0^{\infty} W_n(E) f(E) N(E) dE, \quad (6.215)$$

где $W_n(E)$ — вероятность ионизации, рассчитываемая квантово-механическим путем; $f(E)$ — неравновесная функция распределения; $N(E)$ — плотность квантовых состояний; v — скорость дрейфа.

Для полупроводников с большой диэлектрической проницаемостью вероятность ионизации (для Ge и Si):

$$W_n(E) \approx A(E - E_n)^2, \quad (6.216)$$

где A — коэффициент пропорциональности, а E_n — энергия ионизации.

При ударной ионизации участвуют три носителя заряда, которые примерно поровну распределяют между собой энергию первоначального носителя заряда; пороговая энергия ионизации $E_n = -\frac{3}{2} E_g$, где E_g — ширина запрещенной зоны. Зависимость коэффициента ударной ионизации $\alpha(E)$ от поля в основном определяется видом функции распределения $f(E)$. Как показывает детальный анализ с помощью кинетического уравнения, полное аналитическое выражение для функции распределения в области энергии

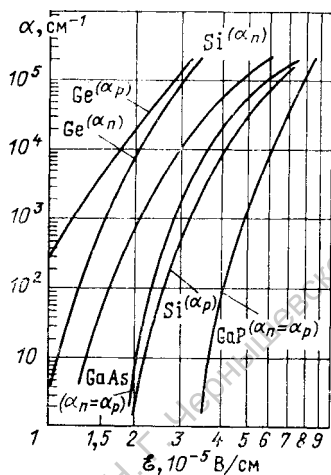


Рис. 6.23. Зависимость коэффициента ударной ионизации от электрического поля для германия, кремния, арсенида галлия и фосфида галлия

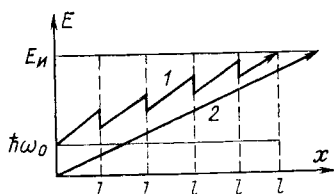


Рис. 6.24. Условная диаграмма ударной ионизации

теряемая электроном, равна $\hbar\omega_0 = \text{const}$, а средняя длина свободного пробега не зависит от энергии, т. е. $l = \text{const}$. Ускорение до порога ионизации $E_{и}$ можно проиллюстрировать следующей условной диаграммой, изображенной на рис. 6.24. Здесь по вертикали отложена энергия, приобретаемая электроном во внешнем электрическом поле, а по горизонтали — расстояние, проходимое электроном в среднем между двумя ионизирующими столкновениями. Как следует из рис. 6.24, существует две возможности ускорения электронов до порога ионизации. В первом случае в очень сильных электрических полях, когда выполняется условие $e\mathcal{E}l \gg \hbar\omega_0$, электроны могут набрать энергию в результате многократных столкновений с фононами, так как относительная потеря энергии в каждом таком столкновении мала (линия 1). В этом случае, как показывает теория,

$$\alpha(E) \sim e^{-E_{и}/\mathcal{E}^2}. \quad (6.217)$$

Здесь можно провести аналогию с диффузионным движением частицы, поэтому указанное приближение для $\alpha(E)$ называют диффузионным.

Во втором случае, когда поля относительно слабые ($e\mathcal{E}l \ll \hbar\omega_0$), носители заряда набирают энергию на расстоянии, равном многим длинам свободного пробега (линия 2). В такой ситуации существует только одна возможность набрать энергию $E_{и}$, когда носители заряда не сталкиваются с фононами до порога ионизации. Поскольку путь до порога ионизации $\mathcal{L}_{и} = E_{и}/e\mathcal{E}$, а средний свободный пробег между столкновениями с фононами равен l , то вероятность пробега до порога ионизации без столкновений есть $e^{-\mathcal{L}_{и}/l} \approx e^{-E_{и}/e\mathcal{E}l}$. В связи с этим функция распределения, а следовательно, и коэффициент ударной ионизации пропорциональны

$$\alpha(E) \sim e^{-E_{и}/e\mathcal{E}l}. \quad (6.218)$$

Функция распределения будет сильно вытянута вдоль направления электрического поля, т. е. в направлении дрейфа носителей, поэтому это приближение для $\alpha(E)$ называют дрейфовым.

Как показали эксперименты, в Ge и Si в основном $\alpha(E)$ определяется законом вида (6.218). Однако в узких *p-n* переходах (особенно четко это проявляется в Si) наблюдается зависимость вида (6.217). Это имеет место при напряженности поля $\mathcal{E} \geq 5 \cdot 10^5$ В/см.

6.12. ТУННЕЛЬНЫЙ ЭФФЕКТ И ЭЛЕКТРОСТАТИЧЕСКАЯ ИОНИЗАЦИЯ

В очень сильных электрических полях становится возможным еще один механизм образования свободных носителей заряда — туннельный эффект, или эффект Зинера. Как известно, у полупроводника, находящегося во внешнем электрическом поле, имеет место наклон энергетических зон (см. § 2.10). Наклон зон тем больше, чем больше напряженность электрического поля $\mathcal{E}$. На рис. 6.25 представлена зонная структура собственного полупроводника при наличии сильного электрического поля. В этом случае возможен переход электрона из валентной зоны в зону проводимости благодаря туннельному эффекту. Вероятность туннельного перехода зависит от высоты и ширины потенциального барьера. В данном случае высота потенциального барьера AB представляет собой ширину запрещенной зоны E_g . Эффективная ширина барьера $AB = \Delta x$ может быть определена из разности потенциальной энергии электрона в зоне проводимости в точке B и в валентной зоне — в точке A . Так как потенциальная энергия с точностью до аддитивной постоянной есть

$$U(B) - U(A) = -e\mathcal{E} \Delta x = -E_g, \quad (6.219)$$

то эффективная ширина барьера

$$\Delta x = \frac{E_g}{e\mathcal{E}}. \quad (6.220)$$

Следовательно, ширина потенциального барьера зависит от напряженности электрического поля.

Переход электрона из точки A в точку B связан с переходом сквозь треугольный потенциальный барьер ABV . Но, как известно из квантовой механики, вероятность перехода для барьера треугольной формы имеет вид:

$$D = D_0 e^{-\frac{8\pi \sqrt{2m_n^*} E_g^{3/2}}{3eh\mathcal{E}}}. \quad (6.221)$$

Из этого выражения следует, что вероятность туннельного перехода из зоны в зону растет экспоненциально с ростом напряженности электрического поля $\mathcal{E}$. Кроме того, вероятность туннельного перехода будет больше у полупроводников с малыми значениями ширины запрещенной зоны E_g и эффективной массы m^* (рис. 6.26).

Вероятность туннельного перехода одна и та же для переходов как из валентной зоны в зону проводимости, так и из зоны проводимости в валентную зону. Но поскольку в валентной зоне электронов намного больше, чем в зоне проводимости, то будет иметь место преимущественный переход электронов из валентной

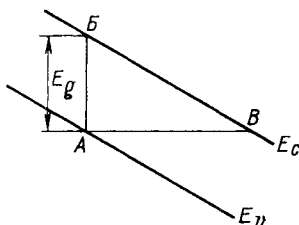


Рис. 6.25. Зонная структура собственного полупроводника при наличии сильного электрического поля

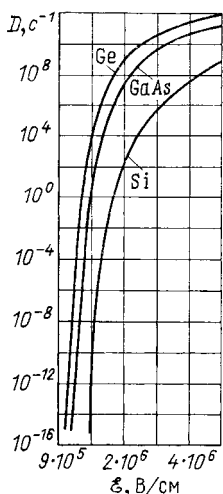


Рис. 6.26. Вероятность туннельного перехода электронов в зависимости от напряженности электрического поля в p - n переходе

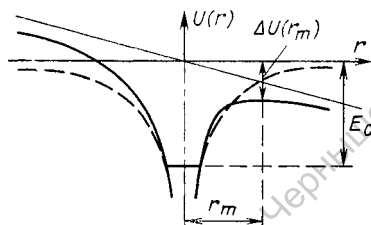


Рис. 6.27. Понижение потенциального барьера под действием внешнего электрического поля

зоны в зону проводимости. Поэтому туннельный эффект может приводить к значительному увеличению концентрации свободных носителей заряда. Он наблюдается в полях 10^6 В/см.

Увеличение концентрации свободных носителей заряда может быть также обусловлено уменьшением энергии ионизации атомов донорной примеси в сильном электрическом поле. На рис. 6.27 представлена потенциальная энергия электрона в атоме донорной примеси без поля (пунктирная кривая) и с внешним электрическим полем (сплошная кривая). При отсутствии электрического поля электрон локализован в потенциальной яме, образованной положительно заряженным донором, а энергия, необходимая для его ионизации, равна E_d .

Внешнее электрическое поле понижает потенциальный барьер в направлении, противоположном направлению электрического поля.

Потенциальная энергия электрона с учетом действия внешнего электрического поля, направленного вдоль вектора $\mathbf{r}$, может быть записана в виде:

$$U(\mathbf{r}) = -e\mathcal{E}r - \frac{e^2}{4\pi\epsilon_r\epsilon_0 r}. \quad (6.222)$$

Первое слагаемое в этом равенстве отражает изменение потенциальной энергии во внешнем электрическом поле напряженностью $\mathcal{E}$, а второе — в кулоновском поле донора. Как следует из (6.222), энергия $U(\mathbf{r})$ в некоторой точке $\mathbf{r}_m$ принимает экстремальное значение, а потенциальный барьер в этой точке снижается на

$$\Delta U(\mathbf{r}_m) = -e\sqrt{e\mathcal{E}/\pi\epsilon_r\epsilon_0}, \quad (6.223)$$

а следовательно, на такое же значение уменьшается и энергия ионизации донора.

Уменьшение энергии ионизации донора увеличивает вероятность термического возбуждения согласно статистике Больцмана на

$$e^{-\Delta U/kT} = e^{-\frac{(e^2/\pi\epsilon_0\epsilon_0)^{1/2}}{kT}} \sqrt{\mathcal{E}}. \quad (6.224)$$

Напряженность электрического поля, при которой имеет место заметное возрастание концентрации электронов, обычно составляет 10^4 — 10^5 В/см. Этот механизм образования свободных носителей заряда проявляется только в том случае, когда концентрация неионизированной донорной примеси значительно превышает концентрацию свободных электронов.

ГЛАВА СЕДЬМАЯ

ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ ЭЛЕКТРОНОВ И ДЫРОК

7.1. РАВНОВЕСНЫЕ И НЕРАВНОВЕСНЫЕ НОСИТЕЛИ ЗАРЯДА

Рассмотрим донорный полупроводник, находящийся в состоянии термодинамического равновесия. В результате тепловой генерации электроны с донорной примеси переходят в зону проводимости. При достаточно высоких температурах будет существенным переброс электронов непосредственно из валентной зоны в зону проводимости. На рис. 7.1 стрелками показаны переходы электронов при тепловом возбуждении. Свободные носители зарядов, возникающие в результате тепловой генерации и находящиеся в термодинамическом равновесии с кристаллической решеткой, называют равновесными.

На рис. 7.1 изображены плотность квантовых состояний $N(E)$ в зоне проводимости и валентной зоне, функция распределения Ферми—Дирака $f_0(E)$ и концентрация равновесных электронов n_0 и дырок p_0 , которые занимают состояния вблизи краев соответствующих зон (заштрихованные области).

Одновременно с генерацией свободных носителей заряда идет процесс рекомбинации: электроны возвращаются в свободные состояния в валентной зоне, в результате чего исчезают свободный электрон и свободная дырка. В условиях термодинамического равновесия процессы генерации и рекомбинации

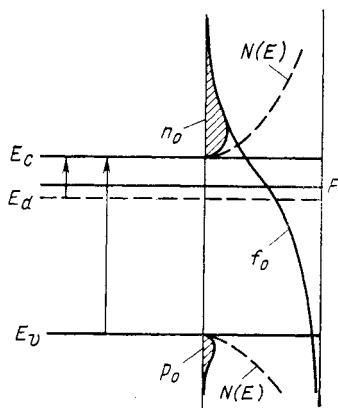


Рис. 7.1. Тепловая генерация свободных носителей заряда и их концентрация в состоянии термодинамического равновесия

взаимно уравниваются. Обозначим через G_0 число генерируемых, а через R_0 — число рекомбинирующих электронно-дырочных пар в единице объема кристалла за одну секунду. Вероятность рекомбинации пропорциональна произведению концентраций свободных носителей заряда, поэтому

$$R_0 = \gamma_r n_0 p_0, \quad (7.1)$$

где γ_r — коэффициент пропорциональности, называемый коэффициентом рекомбинации.

Для равновесного состояния полупроводника справедливо равенство

$$G_0 = R_0, \quad (7.2)$$

которое является выражением принципа детального равновесия.

Помимо тепловой генерации имеются другие механизмы, приводящие к возникновению свободных носителей заряда. Например, они могут образоваться в результате облучения полупроводника светом, разрыва валентных связей в сильных электрических полях или инжекции с помощью p - n перехода. Во всех этих случаях возникает некоторая концентрация электронов n и дырок p , которая отличается от термодинамически равновесной. *Подвижные носители заряда, не находящиеся в термодинамическом равновесии как по концентрации, так и по энергетическому распределению, являются неравновесными носителями заряда, а их концентрацию n , p называют неравновесной. Избыток неравновесной концентрации носителей заряда Δn , Δp в полупроводнике по сравнению с равновесной n_0 , p_0 принято называть избыточной концентрацией носителей заряда.*

При оптической генерации носители заряда могут иметь кинетическую энергию, значительно превышающую среднюю тепловую энергию равновесных частиц. В результате рассеяния на дефектах кристаллической решетки носители заряда довольно быстро передают ей свою избыточную энергию. Допустим, что энергия электронов, генерированных светом, в зоне проводимости составляла 1 эВ. Рассеиваясь на длинноволновых акустических фононах электроны за столкновение будут передавать решетке энергию $2 \cdot 10^{-3}$ эВ (см. с. 151). Таким образом, для рассеяния своей избыточной энергии избыточные электроны должны совершить около 500 столкновений с акустическими фононами. Так как средняя длина свободного пробега электронов имеет порядок 10^{-6} см, а их тепловая скорость при комнатной температуре составляет приблизительно 10^7 см/с и, следовательно, среднее время между двумя столкновениями $\tau \approx 10^{-13}$ с, то через $5 \cdot 10^{-11}$ с избыточные электроны приобретут температуру кристаллической решетки. Поэтому распределение по энергиям неравновесных и равновесных носителей заряда будет одинаковым. Этот процесс сводится к тому, что неравновесные электроны, рассеивая свою избыточную энергию, как бы «опускаются» к нижнему краю зоны проводимости, а неравновесные дырки, рассеивая избыточную энергию, «поднимаются» к верхнему краю валентной зоны. Если при этом концентрация неравновесных

носителей заряда мало отличается от равновесной, то передача кристаллу кинетической энергии неравновесными носителями заряда практически не изменит энергию решетки, а следовательно, и температуру кристалла, поэтому не изменится концентрация равновесных носителей заряда. В этом случае общее количество электронов и дырок будет равно соответственно:

$$\left. \begin{aligned} n &= n_0 + \Delta n; \\ p &= p_0 + \Delta p. \end{aligned} \right\} \quad (7.3)$$

Стационарную концентрацию неравновесных электронов n можно выразить с помощью формулы, аналогичной (4.40):

$$n = n_0 + \Delta n = \int_{E_c}^{\infty} f_n(E) N(E) dE, \quad (7.4)$$

где $f_n = (e^{(E - F_n)/kT} + 1)^{-1}$ — вероятность нахождения электрона на уровне энергии E в зоне проводимости. Поскольку f_n по форме совпадает с распределением Ферми—Дирака (4.28), то по определению F_n есть квазиуровень Ферми для электронов. В соответствии с (4.42) неравновесная концентрация электронов

$$n = N_c F_{1/2}(\eta_n), \quad (7.5)$$

где

$$\eta_n = \frac{F_n - E_c}{kT}. \quad (7.6)$$

есть приведенный квазиуровень Ферми для электронов.

Концентрацию неравновесных электронов в невырожденном полупроводнике с учетом (4.65), (4.66) и (4.69) можно представить в виде

$$n = N_c e^{\eta_n} = n_0 e^{\eta_n - \eta_i} = n_i e^{\eta_n - \eta_i}. \quad (7.7)$$

Аналогично концентрации неравновесных дырок можно выразить следующим образом:

$$p = p_0 + \Delta p = \int_{-\infty}^{E_v} f_p(E) N(E) dE = N_v F_{1/2}(-\eta_p - \varepsilon_i). \quad (7.8)$$

Здесь $f_p = (e^{(E - F_p)/kT} + 1)^{-1}$ — вероятность нахождения дырки на уровне E в валентной зоне, где по определению F_p есть квазиуровень Ферми для дырок, а

$$\eta_p = \frac{F_p - E_v}{kT} \quad (7.9)$$

— приведенный квазиуровень Ферми для дырок.

При отсутствии вырождения, учитывая формулы (4.67) и (4.69), можно написать:

$$p = N_v e^{-\eta_p - \varepsilon_i} = p_0 e^{\eta_n - \eta_p} = n_i e^{\eta_n - \eta_p}. \quad (7.10)$$

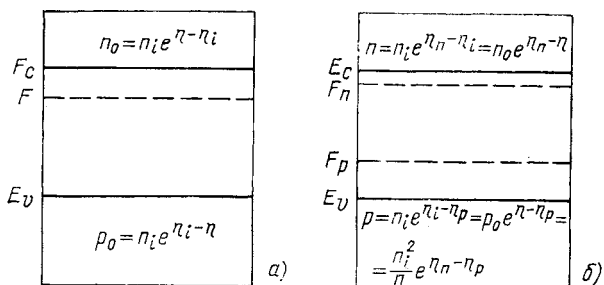


Рис. 7.2. Расщепление уровня Ферми на два квазиуровня — для электронов F_n и дырок F_p
 а — равновесное состояние; б — неравновесное состояние

Таким образом, в неравновесном состоянии уровень Ферми как бы расщепляется на два квазиуровня — для электронов F_n и для дырок F_p , каждый из которых смещен по направлению к своей зоне (рис. 7.2).

Произведение концентраций электронов и дырок для неравновесного состояния отличается теперь от его значения для равновесного состояния:

$$np = n_0 p_0 e^{\eta_n - \eta_p} = n_i^2 e^{\eta_n - \eta_p}. \quad (7.11)$$

Расстояние между приведенными квазиуровнями Ферми ($\eta_n - \eta_p$) характеризует отклонение системы от состояния термодинамического равновесия:

$$np/n_i p_0 = np/n_i^2 = e^{\eta_n - \eta_p}. \quad (7.12)$$

Это соотношение выражает связь между концентрациями электронов и дырок в неравновесном состоянии. При этом чем сильнее различаются квазиуровни Ферми электронов и дырок, тем сильнее отличается произведение концентраций неравновесных носителей заряда от произведения равновесных концентраций.

7.2. БИПОЛЯРНАЯ ОПТИЧЕСКАЯ ГЕНЕРАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА

Рассмотрим полупроводник, в котором оптической генерацией создается некоторая концентрация электронов и дырок. Генерацию носителей заряда, при которой в результате возбуждения возникают пары — электрон и дырка, будем называть биполярной генерацией. В этом случае поглощение кванта света сопровождается разрывом ковалентной связи и количество генерируемых избыточных электронов и дырок одинаково:

$$\Delta n = \Delta p. \quad (7.13)$$

Параллельно с процессом генерации идет процесс рекомбинации, и в стационарном состоянии число электронов и дырок, возбуждаемых за единицу времени в результате разрыва ковалентных связей,

равно числу электронов и дырок, рекомбинирующих за это же время. Поскольку неравновесные носители заряда через очень малое время становятся физически неотличимыми по энергии от равновесных, можно считать, что они имеют тот же коэффициент рекомбинации γ_r , что и равновесные носители заряда.

После выключения возбуждающего света концентрации электронов и дырок уменьшаются в результате рекомбинации. При этом скорость убывания числа свободных электронов и дырок определяется разностью скоростей рекомбинации и тепловой генерации

$$-\left(\frac{dn}{dt}\right)_r = -\left(\frac{dp}{dt}\right)_r = \gamma_r np - G_0. \quad (7.14)$$

Принимая во внимание (7.1), (7.2), (7.3) и (7.13), равенство (7.14) можно переписать следующим образом:

$$-\left(\frac{dn}{dt}\right)_r = \gamma_r (np - n_0 p_0) = \gamma_r (n_0 + p_0 + \Delta n) \Delta n. \quad (7.15)$$

Учитывая, что в случае малого уровня возбуждения $\Delta n \ll (n_0 + p_0)$, получаем:

$$\left(\frac{dn}{dt}\right)_r = -\gamma_r (n_0 + p_0) \Delta n = -\gamma_r (n_0 + p_0) (n - n_0). \quad (7.16)$$

Введем обозначение

$$\tau = \frac{1}{\gamma_r (n_0 + p_0)}. \quad (7.17)$$

Тогда (7.16) примет вид:

$$\left(\frac{dn}{dt}\right)_r = -\frac{n - n_0}{\tau} = -\frac{\Delta n}{\tau}, \quad (7.18)$$

откуда

$$\Delta n = \Delta n(0) e^{-t/\tau}, \quad (7.19)$$

где $\Delta n(0)$ — избыточная концентрация электронов в момент выключения возбуждающего света.

Таким образом, в случае малого уровня биполярной генерации избыточная концентрация электронов и дырок после прекращения возбуждения уменьшается по экспоненциальному закону и за время τ количество их в результате рекомбинации убывает в e раз. Следовательно, τ представляет собой среднее время существования избыточной концентрации электронов и дырок, а поэтому его называют временем жизни неравновесных носителей заряда. Для собственного полупроводника скорости убывания числа электронов и дырок равны и значение τ определяет время жизни электронно-дырочных пар. Значение объемного времени жизни неравновесных носителей заряда в зависимости от типа полупроводникового материала и от степени его чистоты может изменяться в очень широких пределах — от 10^{-2} до 10^{-8} с.

В случае большого уровня биполярного возбуждения, когда $\Delta n \gg (n_0 + p_0)$, из (7.15) найдем, что

$$-\left(\frac{dn}{dt}\right)_r = \gamma_r (\Delta n)^2, \quad (7.20)$$

т. е. скорость рекомбинации зависит от Δn по квадратичному закону. Разделяя переменные, запишем:

$$\frac{dn}{(\Delta n)^2} = -\gamma_r dt, \quad (7.21)$$

откуда

$$\Delta n = \frac{\Delta n(0)}{1 + \gamma_r \Delta n(0) t}. \quad (7.22)$$

Из (7.22) следует, что при квадратичной рекомбинации избыточная концентрация носителей заряда уменьшается по гиперболическому закону.

Если ввести мгновенное время жизни $\tau_{\text{МГН}}$, то в случае квадратичной рекомбинации согласно (7.20)

$$\left(\frac{dn}{dt} \right)_r = -\frac{\Delta n}{\tau_{\text{МГН}}} \quad (7.23)$$

и

$$\tau_{\text{МГН}} = \frac{1}{\gamma_r \Delta n}, \quad (7.24)$$

т. е. $\tau_{\text{МГН}}$ зависит от концентрации избыточных электронов и в связи с этим является переменной величиной. Однако, в каждый момент времени $\tau_{\text{МГН}}$ имеет определенный смысл, являясь мгновенным временем жизни носителей заряда. В силу этого при большом уровне инжекции имеем дело с мгновенным временем жизни, которое согласно (7.23) равно:

$$\tau_{\text{МГН}} = -\frac{\Delta n}{\frac{d \Delta n}{dt}}. \quad (7.25)$$

7.3. МОНОПОЛЯРНАЯ ОПТИЧЕСКАЯ ГЕНЕРАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА. МАКСВЕЛЛОВСКОЕ ВРЕМЯ РЕЛАКСАЦИИ

Процессы генерации неравновесных носителей заряда в примесных полупроводниках могут существенно отличаться от явлений, происходящих в материалах с чисто собственной проводимостью. Рассмотрим донорный полупроводник, в котором облучение светом приводит к перебросу электронов с донорных уровней в зону проводимости. Такой процесс возбуждения носителей заряда будем называть монополярной оптической генерацией. Она характеризуется образованием избыточной концентрации основных носителей заряда (в данном случае электронов). При этом электронейтральность полупроводника не нарушается, так как избыточный заряд свободных электронов, генерируемых светом, скомпенсирован зарядом образовавшихся положительных ионов донорной примеси.

Если неравновесные электроны генерируются в некоторой области образца, то концентрация электронов в этой области будет повышенной: $n = n_0 + \Delta n$. Наличие избыточной концентрации электронов вызовет перемещение их в неосвещенную область полупро-

водника. Поэтому в неосвещенной области образца, куда продиффундируют электроны, возникнет отрицательный объемный заряд избыточных электронов, а в области, откуда уйдут электроны, будет сосредоточен положительный заряд, обусловленный ионами донорной примеси. Если в момент времени $t = 0$ генерация электронов прекратится, то электрическое поле $\mathcal{E}$, порожденное объемными зарядами, вызовет ток проводимости, который в течение некоторого времени приведет к уничтожению объемного заряда.

Изменение плотности пространственного заряда ρ в результате протекания тока, плотность которого равна J , как известно, подчиняется уравнению непрерывности электрического заряда

$$\operatorname{div} J = \operatorname{div} (\sigma \mathcal{E}) = - \frac{\partial \rho}{\partial t}, \quad (7.26)$$

а напряженность электрического поля и объемный заряд связаны уравнением Пуассона

$$\operatorname{div} \mathcal{E} = \rho / \epsilon_r \epsilon_0, \quad (7.27)$$

где ϵ_r — относительная диэлектрическая проницаемость полупроводника; ϵ_0 — электрическая постоянная.

Из уравнений (7.26) и (7.27) следует, что

$$\frac{\partial \rho}{\partial t} = \frac{\sigma}{\epsilon_r \epsilon_0} \rho. \quad (7.28)$$

Отсюда найдем закон изменения объемного заряда во времени

$$\rho = \rho_0 e^{-t/\tau_\mu}, \quad (7.29)$$

где ρ_0 — плотность объемного заряда в момент $t = 0$;

$$\tau_\mu = \epsilon_r \epsilon_0 / \sigma \quad (7.30)$$

есть диэлектрическое, или максвелловское, время релаксации.

Оценим численное значение максвелловского времени релаксации. Для германия, имеющего $\sigma = 1 \text{ Ом}^{-1} \cdot \text{см}^{-1} = 10^2 \text{ См/м}$ и $\epsilon_r = 16$, $\tau_\mu = 10^{-12} \text{ с}$.

Как следует из (7.29), в случае монополярной оптической генерации носителей заряда возникает объемный заряд, который после выключения возбуждающего света со временем уменьшается по экспоненциальному закону с постоянной времени τ_μ . Другими словами, объемный заряд, созданный избыточными основными носителями заряда, в результате тока проводимости после выключения возбуждающего света исчезнет в среднем через τ_μ .

7.4. МЕХАНИЗМЫ РЕКОМБИНАЦИИ

Процессы рекомбинации неравновесных носителей заряда в полупроводниках можно разделить на два основных типа: межзонная рекомбинация и рекомбинация через ловушки.

Межзонная рекомбинация осуществляется при переходе электрона из состояния в зоне проводимости в пустое состояние валентной зоны, что равносильно уничтожению свободного электрона и свободной дырки. Этот процесс совершается при соблюдении законов сохранения энергии и квазимпульса. Если энергия и квазимпульс электрона в зоне проводимости (до рекомбинации) были E' и p' , а после рекомбинации равны соответственно E и p , то законы сохранения энергии и квазимпульса запишутся в виде

$$E' = E + \Delta E; \quad p' = p + Q. \quad (7.31)$$

Здесь Q — квазимпульс, передаваемый электроном кристаллической решетке, а ΔE — энергия, которая выделяется при процессе рекомбинации.

В зависимости от того, каким образом рассеивается избыток энергии ΔE , межзонная рекомбинация делится на три типа:

если энергия, освобождающаяся в процессе рекомбинации, излучается в виде кванта света, рекомбинацию называют излучательной или фотонной;

если энергия, освобождающаяся в процессе рекомбинации, передается решетке, т. е. расходуется на образование фононов, рекомбинация носит название безызлучательной или фононной;

если в акте рекомбинации выделяющаяся энергия передается третьему свободному носителю заряда, рекомбинацию называют ударной, или Оже-рекомбинацией.

В полупроводниковом кристалле всегда имеются дефекты, энергетические уровни которых находятся в запрещенной зоне. Поэтому наряду с межзонной рекомбинацией может идти процесс рекомбинации через локальные уровни энергии. Дефект решетки, способный захватить электрон из зоны проводимости и дырку из валентной зоны, осуществляя их рекомбинацию, называют рекомбинационной ловушкой. В этом случае исчезновение пары — электрона и дырки проводимости — осуществляется следующим образом. Нейтральная рекомбинационная ловушка захватывает электрон из зоны проводимости, который затем через некоторое время перейдет в валентную зону. В зависимости от способа расходования энергии ΔE при переходе электрона на более низкий энергетический уровень рекомбинация через ловушки также может быть фотонной или фононной.

7.5. МЕЖЗОННАЯ ИЗЛУЧАТЕЛЬНАЯ РЕКОМБИНАЦИЯ

При межзонной излучательной рекомбинации электрон из зоны проводимости переходит в валентную зону, при этом излучается квант света, энергия которого равна ширине запрещенной зоны:

$$h\nu = E_g. \quad (7.32)$$

При термодинамическом равновесии количество носителей заряда, рекомбинирующих в единице объема за единицу времени с испусканием фотонов, в соответствии с принципом детального

равновесия равно числу носителей заряда, возбуждаемых при поглощении фотонов. Иными словами, плотность равновесного излучения R_n^0 равна числу поглощаемых квантов R_n^0 в единице объема за одну секунду. Но интенсивность процесса излучательной рекомбинации пропорциональна произведению концентраций равновесных электронов и дырок:

$$R_n^0 = R_p^0 = \gamma_r n_0 p_0 = \gamma_r n_i^2, \quad (7.33)$$

где γ_r — коэффициент межзонной излучательной рекомбинации.

При отклонении от равновесного состояния концентрации неравновесных носителей заряда равны $n = n_0 + \Delta n$ и $p = p_0 + \Delta p$. Они, как отмечалось в § 7.2, характеризуются тем же коэффициентом рекомбинации, что и равновесные. При этом, если не наступает вырождения носителей заряда, сохраняется пропорциональность скорости рекомбинации R произведению концентраций неравновесных электронов и дырок. Поэтому, определяя γ_r из соотношения (7.33), можно написать:

$$R = \gamma_r n p = R_n^0 \frac{np}{n_i^2}. \quad (7.34)$$

Определим время жизни неравновесных носителей заряда при межзонной излучательной рекомбинации. Для электронов в соответствии с (7.18) имеем:

$$\tau_{nr} = -\frac{\Delta n}{dn/dt}. \quad (7.35)$$

Когда внешнее возбуждение снято, скорость изменения концентрации свободных электронов определяется разностью интенсивностей рекомбинации R и равновесной генерации R_n^0 . Так как $R_n^0 = R_p^0$, то

$$-\frac{dn}{dt} = R - R_n^0. \quad (7.36)$$

С учетом этого (7.35) запишется в виде

$$\tau_{nr} = \frac{\Delta n}{R - R_n^0} = \frac{\Delta n}{\Delta R}, \quad (7.37)$$

где ΔR — изменение скорости рекомбинации при отклонении системы от равновесного состояния.

Аналогичное равенство можно записать для дырок:

$$\tau_{pr} = \frac{\Delta p}{R - R_p^0} = \frac{\Delta p}{\Delta R}. \quad (7.38)$$

Величину ΔR определим с помощью (7.34) и (7.33):

$$\Delta R = R - R_n^0 = R_n^0 (np - n_i^2)/n_i^2. \quad (7.39)$$

Подставив в (7.39) значения n и p и заменив n_i^2 на $n_0 p_0$, получим:

$$\Delta R = R_n^0 \frac{n_0 \Delta p + p_0 \Delta n + \Delta n \Delta p}{n_0 p_0}. \quad (7.40)$$

Теперь (7.37) принимает вид:

$$\tau_{nr} = \Delta n \frac{n_0 p_0}{R_n^0 (n_0 \Delta p + p_0 \Delta n + \Delta n \Delta p)}. \quad (7.41)$$

Если $\Delta n = \Delta p$, то (7.41) упрощается:

$$\tau_{nr} = \tau_{pr} = \frac{n_0 p_0}{R_n^0 (n_0 + p_0 + \Delta n)}. \quad (7.42)$$

Проведем анализ полученного выражения. В случае малого уровня возбуждения $\Delta n \ll (n_0 + p_0)$ с учетом (7.33) имеем:

$$\tau_{nr} = \tau_{pr} = \frac{n_i^2}{R_n^0 (n_0 + p_0)} = \frac{1}{\gamma_r (n_0 + p_0)}. \quad (7.43)$$

Для собственного полупроводника, у которого $n_0 = p_0 = n_i$,

$$\tau_{ir} = \frac{n_i}{2R_n^0} = \frac{1}{2\gamma_r n_i} = \frac{e^{E_g/2kT}}{2\gamma_r (N_c N_v)^{1/2}}. \quad (7.44)$$

Для материала n -типа ($n_0 \gg p_0$)

$$\tau_r^n = \frac{p_0}{R_n^0} = 2 \frac{n_i}{n_0} \tau_{ir} = \frac{1}{\gamma_r n_0}, \quad (7.45)$$

а для дырочного полупроводника ($p_0 \gg n_0$)

$$\tau_r^p = \frac{n_0}{R_n^0} = 2 \frac{n_i}{p_0} \tau_{ir} = \frac{1}{\gamma_r p_0}. \quad (7.46)$$

Анализ формул (7.44)–(7.46) показывает, что время жизни неравновесных носителей зарядов в собственном полупроводнике при межзонной излучательной рекомбинации тем меньше, чем выше температура и чем меньше ширина запрещенной зоны (см. табл. 7.1). В примесном полупроводнике время жизни неравновесных носителей заряда меньше, чем τ_{ir} в собственном полупроводнике, и с ростом степени легирования и температуры оно уменьшается.

Зависимость времени жизни для излучательной межзонной рекомбинации от степени легирования, т. е. концентрации носителей заряда при заданной температуре в случае малого уровня возбуждения представлена на рис. 7.3. На этом рисунке логарифмическую шкалу концентрации носителей заряда можно рассматривать так же, как линейную шкалу положения уровня Ферми, так как

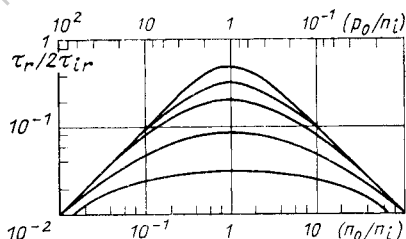


Рис. 7.3. Зависимость времени жизни для межзонной излучательной рекомбинации от концентрации носителей заряда при заданной температуре в случае низкого уровня возбуждения. В максимумах отношение $\Delta n/n_i = 0$ для верхней кривой и соответственно равно 1, 3, 10, 30 для последующих кривых

$n = N_c e^{-(E_c - F)/kT}$, причем ее средняя точка соответствует значению уровня Ферми для собственного полупроводника. Из приведенных зависимостей следует, что с увеличением уровня возбуждения время жизни в собственном полупроводнике резко снижается, а в примесном изменяется сравнительно слабо.

Рассчитаем $R_{и}^0$, используя для этого теорию Эйнштейна для равновесного излучения. Количество квантовых состояний для фотонов в интервале импульса между P и $P + dP$ составляет:

$$P^2 dP/h^3. \quad (7.47)$$

Чтобы определить количество ячеек, занятых фотонами, эту величину следует умножить на функцию распределения Бозе—Эйнштейна:

$$f_{\Phi} = \frac{1}{e^{h\nu/kT} - 1}. \quad (7.48)$$

Общее число фотонов с импульсом P , излучаемых в телесный угол $d\Omega$, равно:

$$dn_{\Phi} = \frac{2}{h^3} f_{\Phi} P^2 dP d\Omega, \quad (7.49)$$

где коэффициент 2 учитывает различную поляризацию света.

Импульс фотона связан с энергией ($E = h\nu$) следующим образом;

$$P = E/v_{\Phi} = h\nu/v_{\Phi} = \bar{n}h\nu/c, \quad (7.50)$$

где $v_{\Phi} = c/\bar{n}$ — скорость фотона в данной среде; c — скорость света в вакууме; $\bar{n}$ — показатель преломления среды (действительная часть показателя).

Из (7.50) получим:

$$dP = \frac{\bar{n}h}{c} dv, \quad (7.51)$$

что при подстановке в (7.49) после интегрирования по всем углам от 0 до 4π дает:

$$dN = \int_0^{4\pi} \frac{2}{h^3} f_{\Phi} P^2 dP d\Omega = \frac{8\pi}{h^3} f_{\Phi} P^2 dP = \frac{8\pi}{c^3} \frac{\bar{n}^3 v^2 dv}{e^{h\nu/kT} - 1}. \quad (7.52)$$

Если обозначить вероятность поглощения фотона с частотой ν через $g(\nu)$, то для интенсивности поглощения, которая в условиях термодинамического равновесия равна интенсивности излучения, получим следующее выражение:

$$R_{и}^0 = R_{и}^0 = \int_0^{\infty} g(\nu) dN = \frac{8\pi}{c^3} \int_0^{\infty} \frac{\bar{n}^3 g(\nu) v^2 dv}{e^{h\nu/kT} - 1}. \quad (7.53)$$

Время жизни фотона τ_{Φ} есть значение, обратное вероятности его поглощения. Величину τ_{Φ} можно определить, если известны длина свободного пробега l_{Φ} и скорость v_{Φ} . В то же время длина свободного

пробега фотона обратно пропорциональна коэффициенту поглощения $\alpha(\nu)$, т. е. $l_\Phi = 1/\alpha(\nu)$. В результате будем иметь:

$$\tau_\Phi = \frac{1}{g(\nu)} = \frac{l_\Phi}{v_\Phi} = \frac{\bar{n}}{c\alpha(\nu)}. \quad (7.54)$$

Подставляя значение $g(\nu)$ из (7.54) в (7.53), находим:

$$R_n^0 = \gamma_r n_i^2 = \frac{8\pi}{c^2} \int_0^\infty \frac{\bar{n}^2 \alpha(\nu) \nu^3 d\nu}{e^{h\nu/kT} - 1}. \quad (7.55)$$

Значения n_i , R_n^0 и значения времени жизни τ_{ir} для излучательных переходов зона проводимости — валентная зона для некоторых полупроводников приведены в табл. 7.1, где $\tau_{\text{экс}}$ — экспериментально измеренная величина.

Из данных табл. 7.1 следует, что время жизни неравновесных носителей заряда в случае межзонной излучательной рекомбинации снижается с уменьшением ширины запрещенной зоны. Рассчитанные значения τ_{ir} относительно близки к экспериментальным только для полупроводников с малой шириной запрещенной зоны. Для материалов с большой шириной запрещенной зоны наблюдается существенное расхождение расчетных данных с экспериментально измеренными. Это обусловлено тем, что у этих полупроводников велика вероятность безызлучательных переходов.

Т а б л и ц а 7.1. Излучательная рекомбинация при 300 К некоторых полупроводников

Полупроводник	E_g , эВ	n_i , см ⁻³	R_n^0 , см ⁻³ ·с ⁻¹	τ_{ir}	$\tau_{\text{экс}}$, мкс
GaAs	1,43	$4,3 \cdot 10^{12}$	$2,4 \cdot 10^{14}$	0,009 с	0,37
InAs	0,31	$1,6 \cdot 10^{16}$	$5,4 \cdot 10^{19}$	15 мкс	0,24
InSb	0,18	$2,0 \cdot 10^{16}$	$1,6 \cdot 10^{22}$	0,62 мкс	0,12
PbS	0,41	$7,1 \cdot 10^{14}$	$2,4 \cdot 10^{19}$	15 мкс	0,21
PbTe	0,32	$4,0 \cdot 10^{15}$	$8,3 \cdot 10^{20}$	2,4 мкс	0,19
PbSe	0,29	$6,2 \cdot 10^{15}$	$1,5 \cdot 10^{21}$	2 мкс	0,25

Согласно формулам (7.45) и (7.46) при увеличении степени легирования время жизни неравновесных носителей заряда снижается. Например, для сильно легированного германия ($n_0 \approx 10^{19}$ см⁻³) оно составляет:

$$\tau_n^r = 2n_i \tau_{ir} / n_0 \approx 2 \cdot 2,4 \cdot 10^{13} \cdot 0,61 / 10^{19} \approx 3 \cdot 10^{-6} \text{ с} = 3 \text{ мкс.}$$

Казалось бы, что с ростом степени легирования материала должна увеличиваться вероятность межзонной излучательной рекомбинации. Однако с повышением концентрации примеси появляется большое количество дефектов и начинает преобладать рекомбинация через ловушки. Поэтому для исследования процессов межзонной излучательной рекомбинации следует использовать наиболее чистые полупроводниковые материалы.

7.6. МЕЖЗОННАЯ УДАРНАЯ РЕКОМБИНАЦИЯ

Если происходит столкновение одновременно двух свободных электронов и одной дырки или двух дырок и одного свободного электрона, может иметь место рекомбинация двух из них с передачей энергии третьему носителю заряда, который переходит на более высокий энергетический уровень в соответствующей зоне, становясь «горячим».

«Горячий» носитель в результате столкновений передаст свою энергию фононам и придет в равновесное состояние с решеткой. Этот процесс рекомбинации является процессом, обратным ударной ионизации, при котором переход свободного носителя заряда в зоне на более низкий энергетический уровень в результате выделившейся энергии происходит при возникновении пары электрон—дырка. На рис. 7.4 схематично изображены процессы ударной рекомбинации с участием в качестве третьего носителя заряда электрона (рис. 7.4, а) или дырки (рис. 7.4, б).

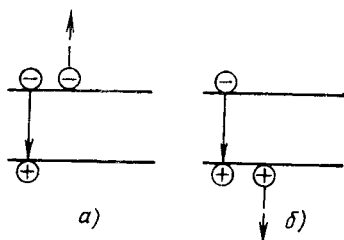


Рис. 7.4. Схема процессов межзональной ударной рекомбинации

Вероятность столкновения пары электрон—дырка со свободным электроном пропорциональна n^2p , а с дыркой — p^2n . Таким образом, уменьшение концентрации свободных электронов или дырок в результате межзональной ударной рекомбинации определится выражением

$$-\frac{dn}{dt} = -\frac{dp}{dt} = \eta^{(n)} n^2 p + \eta^{(p)} p^2 n - \eta^{(n)} n_0^2 p_0 - \eta^{(p)} p_0^2 n_0, \quad (7.56)$$

где $\eta^{(n)}$ и $\eta^{(p)}$ — коэффициенты ударной рекомбинации с участием в качестве третьего носителя заряда электрона или дырки. Два последних члена в (7.56) учитывают интенсивность генерации при термодинамическом равновесии. Следует отметить, что коэффициенты ударной рекомбинации отличаются по размерности от соответствующих коэффициентов в случае рекомбинации другого типа.

Подставив в (7.56) значения n и p , приравняв $\Delta n = \Delta p$ и расположив члены уравнения по степеням Δn , получим:

$$\begin{aligned} -\frac{d}{dt} \Delta n = -\frac{d}{dt} \Delta p = & \Delta n (\eta^{(n)} n_0^2 + 2\eta^{(n)} n_0 p_0 + \\ & + 2\eta^{(p)} n_0 p_0 + \eta^{(p)} p_0^2) + \Delta n^2 (\eta^{(n)} p_0 + 2\eta^{(n)} n_0 + 2\eta^{(p)} p_0 + \eta^{(p)} n_0) + \\ & + \Delta n^3 (\eta^{(n)} + \eta^{(p)}). \end{aligned} \quad (7.57)$$

Находим выражение для времени жизни неравновесных носителей заряда в случае межзональной ударной рекомбинации согласно (7.18), считая, что уровень инжекции мал [$\Delta n \ll (n_0 + p_0)$]. Для

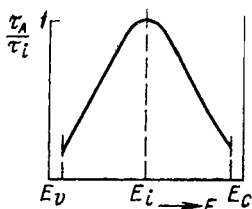


Рис. 7.5. Время жизни при ударной межзонной рекомбинации в зависимости от положения уровня Ферми

этого отбросим в (7.57) члены, содержащие Δn во второй и третьей степени, и будем иметь:

$$\tau_{nA} = \tau_{pA} = - \frac{\Delta n}{\frac{d}{dt} \Delta n} = \frac{1}{\eta^{(n)} n_0^2 + 2\eta^{(n)} n_i^2 + 2\eta^{(p)} n_i^2 + \eta^{(p)} p_0^2}. \quad (7.58)$$

Рассмотрим частные случаи.

1. Собственный полупроводник ($n_0 = p_0 = n_i$)

$$\tau_{iA} = \frac{1}{3(\eta^{(n)} + \eta^{(p)}) n_i^2} = \frac{1}{3(\eta^{(n)} + \eta^{(p)}) N_c N_v} e^{E_g/kT}. \quad (7.59)$$

Как видно, время жизни неравновесных носителей заряда в собственном полупроводнике при межзонной ударной рекомбинации снижается с ростом температуры и с уменьшением ширины запрещенной зоны.

2. Донорный полупроводник ($n_0 \gg p_0$; $n_0 \gg n_i$)

$$\tau_A^n = 1/\eta^{(n)} n_0^2. \quad (7.60)$$

3. Акцепторный полупроводник ($p_0 \gg n_0$; $p_0 \gg n_i$)

$$\tau_A^p = 1/\eta^{(p)} p_0^2. \quad (7.61)$$

Из полученных выражений следует, что время жизни неравновесных носителей заряда для собственного полупроводника при межзонной ударной рекомбинации максимально. В примесном полупроводнике время жизни при этом типе рекомбинации сильно зависит от концентрации основных носителей заряда. При ее увеличении возрастает число столкновений неравновесных электронно-дырочных пар с основными носителями заряда и время жизни снижается. Зависимость τ_A/τ_i от положения уровня Ферми приведена на рис. 7.5.

Экспериментальные исследования показывают, что ударная межзонная рекомбинация наблюдается очень редко. Она становится заметной лишь при довольно высоких температурах в полупроводниках с малой шириной запрещенной зоны (например, в InSb).

При высоком уровне возбуждения можно считать, что $\Delta n \gg (n_0 + p_0)$ и $\Delta p \gg (n_0 + p_0)$. Для этого случая в (7.57) можно отбросить два первых слагаемых, т. е. члены, содержащие Δn и Δn^2 . В результате будем иметь:

$$- \frac{d}{dt} \Delta n = - \frac{d}{dt} \Delta p = \Delta n^3 (\eta^{(n)} + \eta^{(p)}). \quad (7.62)$$

Мгновенное время жизни при этом

$$\tau_{\text{мгн}} = \frac{\Delta n}{\frac{d}{dt} \Delta n} = \frac{1}{(\eta^{(n)} + \eta^{(p)}) \Delta n^2}. \quad (7.63)$$

Из (7.63) следует, что мгновенное время жизни для ударной межзонной рекомбинации при высоком уровне возбуждения очень сильно зависит от концентрации избыточных носителей заряда.

7.7. РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА ЧЕРЕЗ ЛОВУШКИ

Реальные полупроводниковые материалы содержат обычно смеси нескольких типов, каждая из которых может создавать один или несколько уровней в запрещенной зоне полупроводника.

Дефекты решетки, обычно нейтральные в состоянии термодинамического равновесия и способные захватывать подвижные носители заряда одного знака и освобождать их, называют ловушками или центрами захвата. Ограничимся рассмотрением простейшего случая, когда в полупроводнике имеется один тип ловушек, создающих единственный энергетический уровень.

Пусть концентрация ловушек равна N_t , а их энергетический уровень E_t лежит вблизи середины запрещенной зоны. Процесс рекомбинации через эти ловушки проиллюстрирован на рис. 7.6.

В полупроводнике донорного типа, имеющем большое число свободных электронов, рекомбинация электрона и дырки проходит следующим образом. Нейтральная ловушка захватывает электрон из зоны проводимости, приобретая отрицательный заряд (переход 1), а затем электрон с уровня ловушки переходит в валентную зону, что эквивалентно захвату отрицательно заряженной ловушкой дырки (переход 2). При снижении концентрации свободных дырок вероятность второго процесса уменьшается и возможна эмиссия электрона, т. е. обратный переход электрона с ловушки в зону проводимости (переход 3).

Если материал содержит большое количество свободных дырок, они интенсивно захватываются ловушками, т. е. электроны переходят с ловушки в валентную зону (переход 4), а затем происходит захват электрона из зоны проводимости (переход 5). При малой концентрации электронов проводимости возможен процесс освобождения дырки, т. е. возвращение дырки в валентную зону (переход 6).

Перейдем к количественному описанию рассмотренных процессов. Если через $f_t = f(E_t)$ обозначим вероятность заполнения ловушки электроном, то $1 - f_t$ будет вероятностью того, что на ловушке нет электрона, т. е. она свободна. Тогда интенсивность захвата ловушками электронов $R_{\text{зах}}^n$, которая пропорциональна концентрации электронов в зоне проводимости n и числу свободных мест на уровне ловушек $N_t (1 - f_t)$, равна:

$$R_{\text{зах}}^n = \gamma_n n N_t (1 - f_t), \quad (7.64)$$

где γ_n — коэффициент захвата неравновесных электронов ловушками.

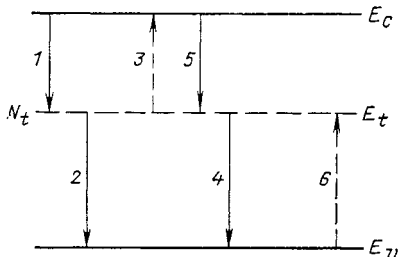


Рис. 7.6. Переходы электронов, связанные с рекомбинацией через ловушки

Количество электронов $R_{\text{воз}}^n$, возвращающихся с ловушек в зону проводимости в результате эмиссии, пропорционально концентрации электронов на ловушках $N_i f_i$:

$$R_{\text{воз}}^n = \beta_n f_i N_i, \quad (7.65)$$

где β_n — коэффициент ионизации электронов с ловушек.

Изменение концентрации неравновесных электронов в зоне проводимости определяется разницей интенсивностей процессов захвата $R_{\text{зах}}^n$ и освобождения $R_{\text{воз}}^n$:

$$-\frac{dn}{dt} = R_{\text{зах}}^n - R_{\text{воз}}^n = \gamma_n n N_i (1 - f_i) - \beta_n N_i f_i. \quad (7.66)$$

Чтобы определить связь между коэффициентами γ_n и β_n , рассмотрим состояние термодинамического равновесия. В состоянии термодинамического равновесия вероятность заполнения уровня ловушки электроном есть

$$f_{0i} = f_0(E_i) = \frac{1}{e^{(E_i - F)/kT} + 1}.$$

Если в неравновесном состоянии не наступает вырождение, то неравновесные носители заряда большую часть своего времени жизни физически неотличимы от равновесных. Поэтому можно считать, что и равновесные носители заряда характеризуются также коэффициентами захвата γ_n и ионизации β_n . На основании (7.66) в случае термодинамического равновесия, для которого $\frac{dn_0}{dt} = 0$, найдем, что

$$\begin{aligned} \beta_n &= \gamma_n n_0 \left(\frac{1}{f_{0i}} - 1 \right) = \gamma_n N_i e^{-(E_c - F)/kT} e^{(E_i - F)/kT} = \\ &= \gamma_n N_i e^{-(E_c - E_i)/kT} = \gamma_n n_i. \end{aligned} \quad (7.67)$$

Здесь

$$n_i = N_i e^{-(E_c - E_i)/kT} \quad (7.68)$$

есть равновесная концентрация электронов в зоне проводимости, когда уровень Ферми совпадает с уровнем ловушки.

Используя (7.67), преобразуем уравнение (7.66) к виду

$$-\frac{dn}{dt} = \gamma_n N_i [n(1 - f_i) - n_i f_i]. \quad (7.69)$$

Это выражение определяет скорость изменения концентрации неравновесных электронов в зоне проводимости в результате процессов захвата электронов ловушками и освобождения их из ловушек в зону проводимости.

Скорость изменения концентрации неравновесных дырок в валентной зоне запишется следующим образом:

$$-\frac{dp}{dt} = \gamma_p \rho N_i f_i - \beta_p N_i (1 - f_i). \quad (7.70)$$

Здесь первое слагаемое определяет количество дырок, захватываемых ловушками, а второе — число дырок, возвращающихся

из ловушек в валентную зону. Величины γ_p и β_p представляют собой соответственно коэффициенты захвата и ионизации для дырок.

Из условия термодинамического равновесия ($\frac{dp_0}{dt} = 0$) найдем:

$$\begin{aligned}\beta_p &= \gamma_p p_0 \frac{f_{0t}}{1 - f_{0t}} = \gamma_p N_v e^{(E_v - F)/kT} e^{-(E_t - F)/kT} = \\ &= \gamma_p N_v e^{(E_v - E_t)/kT} = \gamma_p p_1,\end{aligned}\quad (7.71)$$

где

$$p_1 = N_v e^{(E_v - E_t)/kT} \quad (7.72)$$

есть равновесная концентрация дырок в валентной зоне, когда уровень Ферми совпадает с уровнем ловушек.

Подставляя (7.71) в (7.70), имеем:

$$-\frac{dp}{dt} = \gamma_p N_t [p f_t - p_1 (1 - f_t)]. \quad (7.73)$$

Это выражение определяет скорость изменения концентрации неравновесных дырок в валентной зоне в результате процессов захвата дырок ловушками и освобождения их с ловушек в валентную зону.

Используя (7.69) и (7.73), находим изменение количества электронов, находящихся на ловушках. Оно определяется изменением концентраций неравновесных электронов в зоне проводимости и дырок в валентной зоне:

$$\frac{dn}{dt} - \frac{dp}{dt} = \gamma_p N_t [p f_t - p_1 (1 - f_t)] - \gamma_n N_t [n (1 - f_t) - n_1 f_t]. \quad (7.74)$$

Запишем условие локальной электронейтральности для одномерного полупроводника, содержащего некоторое количество ионов донорной примеси N_d^+ . Для равновесного случая условие электронейтральности имеет вид:

$$n_0 + N_t f_{0t} = p_0 + N_d^+. \quad (7.75)$$

Будем считать, что число ионов донорной примеси не меняется и в неравновесном состоянии. Тогда

$$n + N_t f_t = p + N_d^+. \quad (7.76)$$

Вычитая (7.75) из (7.76), получаем соотношение для избыточной концентрации носителей заряда:

$$n - n_0 + N_t (f_t - f_{0t}) = p - p_0$$

или

$$\Delta n + N_t \Delta f_t = \Delta p. \quad (7.77)$$

Если концентрация ловушек в полупроводнике мала по сравнению с плотностью избыточных носителей заряда, т. е. $N_t \ll \Delta n$, Δp , то количеством электронов $N_t \Delta f_t$, оседающих на ловушках, можно пренебречь. При этом условии избыточная концентрация электронов равна избыточной концентрации дырок, т. е.

$$\Delta n = \Delta p. \quad (7.78)$$

В этом случае $\frac{dn}{dt} = -\frac{dp}{dt}$ и из (7.74), приравняв его нулю, найдем функцию распределения для электронов, находящихся на ловушках:

$$f_t = \frac{\gamma_n n + \gamma_p p_1}{\gamma_n (n + n_1) + \gamma_p (p + p_1)}. \quad (7.79)$$

Подставляя найденное значение f_t в (7.69), получаем:

$$-\frac{dn}{dt} = \frac{\gamma_n \gamma_p N_t (np - n_1 p_1)}{\gamma_n (n + n_1) + \gamma_p (p + p_1)}. \quad (7.80)$$

На основании этого соотношения можно определить время жизни неравновесных носителей заряда при рекомбинации через ловушки. Учитывая, что

$$n_1 p_1 = n_i^2 = n_0 p_0, \quad (7.81)$$

в результате будем иметь:

$$\begin{aligned} \tau = -\frac{\Delta n}{\frac{dn}{dt}} &= \frac{\gamma_n (n_0 + n_1 + \Delta n) + \gamma_p (p_0 + p_1 + \Delta p)}{\gamma_n \gamma_p N_t (n_0 + p_0 + \Delta n)} = \\ &= \frac{1}{\gamma_p N_t} \frac{n_0 + n_1 + \Delta n}{n_0 + p_0 + \Delta n} + \frac{1}{\gamma_n N_t} \frac{p_0 + p_1 + \Delta n}{n_0 + p_0 + \Delta n}. \end{aligned} \quad (7.82)$$

Введем обозначения

$$\tau_{p0} = 1/\gamma_p N_t; \quad \tau_{n0} = 1/\gamma_n N_t. \quad (7.83)$$

С учетом (7.83) формула (7.82) может быть записана в виде

$$\begin{aligned} \tau &= \tau_{p0} (n_0 + n_1 + \Delta n)/(n_0 + p_0 + \Delta n) + \\ &+ \tau_{n0} (p_0 + p_1 + \Delta p)/(n_0 + p_0 + \Delta p). \end{aligned} \quad (7.84)$$

В случае малого уровня возбуждения, когда $\Delta n \ll (n_0 + p_0)$, время жизни неравновесных носителей заряда в зависимости от положения уровня Ферми при постоянной температуре, т. е. от степени легирования материала, будет равно

$$\tau = \tau_{p0} \frac{n_0 + n_1}{n_0 + p_0} + \tau_{n0} \frac{p_0 + p_1}{n_0 + p_0}. \quad (7.85)$$

Итак, при малом уровне возбуждения время жизни электронно-дырочных пар не зависит от концентрации избыточных носителей заряда, а определяется значениями равновесных концентраций электронов и дырок и положением энергетического уровня рекомбинационной ловушки, задающим величины n_1 и p_1 .

На основании формулы (7.85) для зависимости $\ln \tau$ от положения уровня Ферми можно выделить четыре основные области (рис. 7.7).

Область 1. Сильно легированный полупроводник n -типа. Уровень Ферми расположен ниже дна зоны проводимости, но выше энергетического уровня рекомбинационной ловушки, т. е. $E_t < < F < E_c$. В этом случае справедливы следующие соотношения концентраций носителей заряда:

$$n_0 \gg p_0; \quad n_0 \gg n_1; \quad n_0 \gg p_1.$$

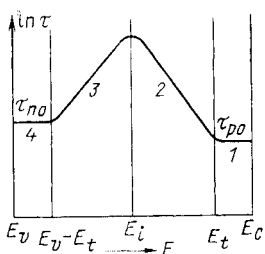


Рис. 7.7. Зависимость времени жизни неравновесных носителей заряда от положения уровня Ферми

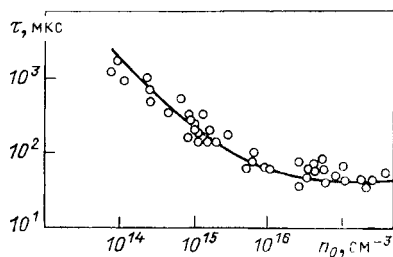


Рис. 7.8. Зависимость времени жизни дырок в электронном германии от равновесной концентрации электронов

Учитывая эти неравенства, из (7.85) получаем:

$$\tau = \tau_{p0} = 1/\gamma_p N_i. \quad (7.86)$$

Таким образом, в сильно легированном полупроводнике n -типа время жизни неравновесных электронов и дырок постоянно и равно $\tau = \tau_{p0}$, т. е. определяется временем жизни неосновных равновесных носителей заряда — дырок. В этом случае скорость процесса рекомбинации определяется числом актов захвата дырок ловушками, занятыми электронами. Действительно, как показывает эксперимент (рис. 7.8), время жизни дырок в сильно легированном электронном германии постоянно.

Проведем анализ (7.86). Это равенство выражает собой тот факт, что каждая дырка в материале n -типа, перемещаясь в валентной зоне, обладает определенной вероятностью быть захваченной отрицательно заряженной рекомбинационной ловушкой. Для количественной оценки этого процесса используем эффективное сечение захвата дырки σ_p . Согласно (5.48) вероятность захвата неравновесного носителя заряда одним центром пропорциональна эффективному сечению захвата ловушкой дырки и тепловой скорости носителей v_0 , а поэтому

$$\gamma_p = \sigma_p v_0. \quad (7.87)$$

Следовательно, время жизни дырки в электронном полупроводнике равно

$$\tau = \tau_{p0} = 1/\gamma_p N_i = 1/\sigma_p N_i v_0. \quad (7.88)$$

Оно уменьшается при увеличении концентрации рекомбинационных ловушек и повышении температуры образца, приводящей к росту v_0 .

Область 2. Полупроводник n -типа легирован слабо, так что уровень Ферми лежит в запрещенной зоне ниже энергетического уровня ловушек, но выше середины запрещенной зоны, т. е. $E_i < F < E_t$. В этом случае имеют место следующие соотношения для концентраций носителей заряда:

$$n_0 \gg p_0; \quad n_0 \gg p_i; \quad n_0 < n_i.$$

Соответственно выражение для τ приобретает вид

$$\tau \approx \tau_{p0} \frac{n_1}{n_0} = \tau_{p0} e^{(E_t - F)/kT}. \quad (7.89)$$

Из (7.89) следует, что по мере понижения уровня Ферми время жизни неравновесных носителей заряда τ растет по экспоненциальному закону. При этом чем ниже опускается уровень Ферми, тем меньше степень заполнения ловушек электронами, поэтому вероятность захвата ловушкой дырки снижается, что ведет к увеличению времени жизни дырки. Это подтверждается данными, представленными на рис. 7.7.

Область 3. В слабо легированном полупроводнике p -типа положение уровня Ферми определяется неравенством $(E_v - E_t) < F < E_i$, что дает следующие соотношения для концентраций носителей заряда:

$$p_0 \gg n_0; p_0 \gg p_1; n_1 > p_0.$$

На основании этого из (7.85) получаем формулу для времени жизни:

$$\tau \approx \tau_{p0} \frac{n_1}{p_0} = \tau_{p0} \frac{N_c}{N_v} e^{-(E_c + E_v - E_t - F)/kT}. \quad (7.90)$$

Видно, что для акцепторного полупроводника по мере понижения уровня Ферми от середины запрещенной зоны в результате легирования время жизни неравновесных носителей заряда уменьшается по экспоненциальному закону. Это происходит потому, что в слабо легированном полупроводнике p -типа почти все ловушки свободные и охотно захватывают электроны из зоны проводимости. С увеличением количества дырок проводимости растет вероятность их рекомбинации с электронами, попавшими на ловушки. При этом интенсивность процесса освобождения электронов с ловушек снижается и время жизни неравновесных носителей заряда уменьшается (рис. 7.7).

Область 4. Сильно легированный материал p -типа, для которого положение уровня Ферми определяется неравенством вида $E_v < F < (E_v - E_t)$.

В этом случае

$$p_0 \gg n_0; p_0 \gg p_1; p_0 \gg n_1,$$

а из (7.85) следует:

$$\tau = \tau_{n0} = 1/\gamma_n N_t = 1/\sigma_n N_t v_0, \quad (7.91)$$

где σ_n — эффективное сечение захвата электрона.

Таким образом, в сильно легированном полупроводнике p -типа время жизни пары электрон—дырка постоянно, равно $\tau = \tau_{n0}$, определяется временем жизни равновесных электронов — неосновных носителей заряда и не зависит от положения уровня Ферми. У такого полупроводника все ловушки свободны от электронов и каждый электрон, захваченный ловушкой, немедленно рекомбинирует с дыркой, так как количество дырок в валентной зоне очень велико. Процесс обратного переброса электронов с ловушек в зону проводимости практически полностью отсутствует и не влияет на время жизни.

Рассмотрим теперь случай большого уровня возбуждения избыточных носителей заряда, когда $\Delta n \gg n_0$; $\Delta p \gg p_0$, а значит, $\Delta n \gg n_1$ и $\Delta p \gg p_1$. Из (7.84) получаем:

$$\tau_{\infty} \approx \tau_{p0} + \tau_{n0} = \frac{\gamma_n + \gamma_p}{\gamma_n \gamma_p N_t}. \quad (7.92)$$

Как следует из этого соотношения, при большом уровне возбуждения время жизни неравновесных электронов и дырок τ_{∞} при рекомбинации через ловушки не зависит от концентраций электронов и дырок, а определяется лишь количеством и свойствами ловушек.

7.8. ТЕМПЕРАТУРНАЯ ЗАВИСИМОСТЬ ВРЕМЕНИ ЖИЗНИ НОСИТЕЛЕЙ ЗАРЯДА ПРИ РЕКОМБИНАЦИИ ЧЕРЕЗ ЛОВУШКИ

Рассмотрим теперь, как меняется время жизни неравновесных носителей заряда при рекомбинации через ловушки в полупроводнике заданного состава при изменении температуры исследования. Этот анализ проведем на примере полупроводника электронного типа. Допустим, что 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 уровнем 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 неравенствам $n_0 \gg n_1$, $n_0 \gg p_0$ и $n_0 \gg p_1$. Следовательно, в этой области в соответствии с (7.86) время жизни неравновесных дырок постоянно и равно τ_{p0} (участок 1' на рис. 7.9, в, на котором представлена температурная зависимость времени жизни).

По мере дальнейшего повышения температуры при переходе к области Б, когда $E_d < F < E_t$, т. е. в области полной ионизации донорной

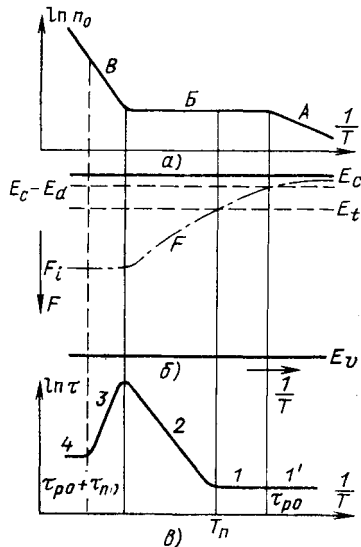


Рис. 7.9. Зависимость равновесной концентрации (а), уровня Ферми (б) и времени носителей жизни заряда (в) от температуры для донорного полупроводника

примеси уровень Ферми понижается и достигает уровня рекомбинационной ловушки. На этом участке концентрация электронов n_0 постоянна, а концентрация $n_1 = N_c e^{-(E_c - E_t)/kT}$ 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 (7.86) время жизни также постоянно 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, в). Здесь по-прежнему $n_0 = \text{const}$, а n_1 значительно увеличивается, поэтому из (7.90) будем иметь:

$$\begin{aligned} \tau &= \tau_{p0} \frac{n_1}{n_0} = \tau_{p0} \frac{N_c}{n_0} e^{-(E_c - E_t)/kT} = \\ &= \frac{\tau_{p0}}{n_0} 2 \left(\frac{2\pi m_n^* k}{h^2} \right)^{3/2} T^{3/2} e^{-(E_c - E_t)/kT}, \end{aligned} \quad (7.93)$$

т. е. время жизни неравновесных носителей заряда 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 уровней 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 (в В-области) с ростом температуры, хотя положение уровня Ферми практически остается постоянным (рис. 7.9, в), концентрация носителей заряда в зонах резко увеличивается, поэтому увеличивается заполнение ловушек носителями заряда и время жизни неравновесных электронов и дырок 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 (7.85) запишется в виде

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

$$= \frac{\tau_{p0}}{2} \left[1 + \left(\frac{N_c}{N_v} \right)^{1/2} e^{-\frac{\frac{E_c + E_v}{2} - E_t}{kT}} \right] + \\ + \frac{\tau_{n0}}{2} \left[1 + \left(\frac{N_i}{N_c} \right)^{1/2} e^{\frac{\frac{E_c + E_v}{2} - E_t}{kT}} \right]. \quad (7.95)$$

Если уровень рекомбинационных ловушек расположен достаточно далеко от середины запрещенной зоны, то в области начала собственной проводимости можно выделить участок, для которого выполняется неравенство $\left(\frac{E_c + E_v}{2} - E_t \right) \gg kT$. Для этого случая (7.95) приобретает вид:

$$\tau \approx \frac{\tau_{n0}}{2} \left(\frac{N_v}{N_c} \right)^{1/2} e^{\frac{\frac{E_c + E_v}{2} - E_t}{kT}}. \quad (7.96)$$

Соответствующий участок на кривой $\ln \tau (1/T)$ характеризуется уменьшением времени жизни неравновесных электронов и дырок (участок 3 на рис. 7.9, в).

При дальнейшем повышении температуры, когда в области собственной проводимости будет иметь место неравенство $\left(\frac{E_c + E_v}{2} - E_t \right) \ll kT$, (7.95) можно записать приближенно следующим образом:

$$\tau \approx \frac{\tau_{p0}}{2} \left(1 + \sqrt{\frac{N_c}{N_v}} \right) + \frac{\tau_{n0}}{2} \left(1 + \sqrt{\frac{N_v}{N_c}} \right). \quad (7.97)$$

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

$$\tau = \tau_{p0} + \tau_{n0} = (\gamma_n + \gamma_p) / \gamma_n \gamma_p N_t. \quad (7.98)$$

Таким образом, когда в области собственной проводимости с увеличением температуры концентрация свободных носителей заряда становится очень высокой, время жизни неравновесных электронов и дырок определяется только параметрами рекомбинационных ловушек и не зависит от температуры (участок 4 на рис. 7.9, в).

Зная концентрацию электронов при насыщении, можно оценить E_t по температуре перехода T_{II} от области 1 к области 2 (рис. 7.9, в). При условии $T = T_{II}$, а $F = E_t$ находим:

$$n_0 = N_c e^{-(E_c - E_t)/kT_{II}}, \quad (7.99)$$

откуда

$$E_c - E_t = kT_{II} \ln(N_c/n_0). \quad (7.100)$$

Экстраполируя прямую область 2 до пересечения с осью ординат, определяем в соответствии с (7.93) значение $\ln(\tau_{p0} N_c/n_0)$, а зная τ_0 из участка 1, находим $\ln(N_c/n_0)$ и подставляем в формулу

(7.100). Следовательно, по снятой экспериментальной зависимости $\ln \tau (1/T)$, подобной изображенной на рис. 7.9, в, можно рассчитать глубину залегания уровня рекомбинационных ловушек E_t .

7.9. ЦЕНТРЫ ЗАХВАТА И РЕКОМБИНАЦИОННЫЕ ЛОВУШКИ

При анализе процессов рекомбинации носителей заряда через ловушки учитывался захват электронов и дырок ловушками и тепловой заброс захваченных 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 интенсивностей 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 имеется состояние, отличное от термодинамического равновесия, концентрации 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 или прилипания электронов. Вблизи верхнего края валентной зоны располагаются уровни центров захвата дырок.

У 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 отношение вероятности захвата носителя заряда противоположного знака к вероятности теплового заброса носителя заряда основ-

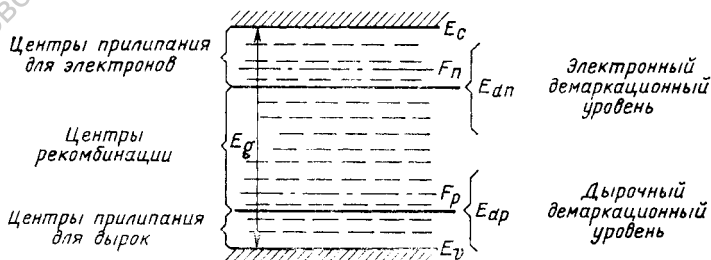


Рис. 7.10. Полупроводник, содержащий центры захвата и рекомбинационные ловушки

ного знака с ловушки в соответствующую энергетическую зону. Применительно к электронам зоны проводимости k_n будет равен отношению вероятности захвата отрицательно заряженной ловушкой дырки к вероятности теплового заброса электрона из ловушки в зону проводимости. Как следует из (7.70), скорость захвата ловушкой дырок определяется значением $\gamma_p N_{ifip}$. Интенсивность процесса теплового возбуждения электронов из ловушек согласно (7.69) равна $\gamma_n N_{ifin}$. Здесь γ_n , γ_p — вероятности захвата электрона и дырки соответственно; f_i — 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:

$$k_n = \frac{\gamma_p p}{\gamma_n n_1} = \frac{\gamma_p p}{\gamma_n N_{ce}^{-(E_c - E_t)/kT}}. \quad (7.101)$$

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

$$\gamma_p N_v e^{(E_v - F_p)/kT} = \gamma_n N_c e^{-(E_c - E_{dn})/kT}. \quad (7.102)$$

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 разность $F_p - E_v$ и электронный демаркационный уровень приближается к дну зоны проводимости. В этом случае будет осуществляться переход центров захвата в рекомбинационные ловушки. С ростом же температуры демаркационный уровень сдвигается к середине запрещенной зоны и рекомбинационные ловушки будут переходить в центры захвата.

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

$$k_p = \frac{\gamma_n n}{\gamma_p p_1}. \quad (7.104)$$

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 центрами захвата электронов. Ловушки, которые имеют уровни энергии ниже дырочного демаркационного уровня, являются центрами захвата дырок. Между ними располагаются энергетические уровни рекомбинационных ловушек (см. рис. 7.10). Необходимо отметить, что такая классификация энергетических уровней ловушек является весьма условной. При ее проведении, в частности, считалось, что отношение коэффициентов захвата γ_n и γ_p изменяется слабо при перемещении уровня ловушек. Центры 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.

ГЛАВА ВОСЬМАЯ

ДИФфуЗИЯ И ДРЕЙФ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА

8.1. УРАВНЕНИЕ НЕПРЕРЫВНОСТИ

Рассмотрим полупроводник, в котором в результате воздействия каких-либо внешних возбуждающих факторов возникли неравновесные носители заряда. В этом случае концентрации неравновесных электронов и дырок являются функциями пространственных координат и времени $n(x, y, z, t)$ и $p(x, y, z, t)$ и скорости их изменения могут быть выражены с помощью уравнения непрерывности.

Вывод уравнения непрерывности проведем для одномерного образца, в котором концентрация электронов изменяется только в направлении оси x , как это представлено на рис. 8.1. Выделим в полупроводнике слой толщиной dx и сечением 1 см^2 . Объем этого слоя будет численно равен dx . Допустим, что в момент времени t концентрация электронов была $n(x, t)$, в объеме dx соответственно $n(x, t) dx$, а в момент времени $t + dt$ будет $n(x, t + dt) dx$ и изменение концентрации электронов за время dt в объеме dx составит:

$$n(x, t + dt) dx - n(x, t) dx = \frac{\partial n}{\partial t} dx dt. \quad (8.1)$$

Это изменение концентрации электронов может происходить в результате процессов генерации, рекомбинации, диффузии и дрейфа носителей заряда. Рассмотрим эти процессы.

1. Под процессами генерации понимаются все механизмы, посредством которых электроны, находящиеся в валентной зоне, на уровнях примеси или уровнях прилипания, могут быть переведены в зону проводимости. Во всех этих случаях нужно учитывать генерацию свободных носителей заряда в результате как тепловых процессов, так и внешних воздействий. К внешним воздействиям относятся действие света, ядерных частиц, неравновесного излучения, испускаемого самим полупроводником в его объеме, и т. п. Обозначим скорость этих двух типов генерации соответственно через G_0 и G . Тогда полная скорость генерации носителей заряда будет равна $G_0 + G$.

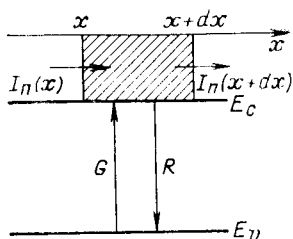


Рис. 8.1. К выводу уравнения непрерывности

Будем считать, что в рассматриваемом случае возбуждение электронов происходит за счет поглощения света полупроводником и в 1 см^3 за 1 с возникает G пар электрон—дырка. Тогда в объеме dx за время dt будет создано электронов в количестве

$$G dx dt. \quad (8.2)$$

2. Изменение концентрации носителей заряда в объеме dx происходит также и в результате рекомбинации. Полную скорость рекомбинации можно представить в виде $R_0 + R$, где R_0 — скорость рекомбинации при тепловом равновесии, R — скорость рекомбинации при наличии внешнего воздействия. В случае линейной рекомбинации, когда время жизни неравновесных электронов τ_n не зависит от концентрации электронов, изменение количества электронов вследствие рекомбинации R в объеме dx за время dt согласно (7.18) составит:

$$R dx dt = - \left(\frac{dn}{dt} \right)_r dx dt = \frac{n - n_0}{\tau_n} dx dt. \quad (8.3)$$

3. Изменение количества носителей заряда в объеме dx может быть обусловлено, кроме того, процессами диффузии и дрейфа. Если $I_n(x, t)$ — поток электронов, проходящих через 1 см^2 поверхности за 1 с , то за время dt через границу слоя x в объем dx войдут электроны в количестве $I_n(x, t) dt$, а через границу $x + dx$ выйдет $I_n(x + dx, t) dt$ электронов. Следовательно, изменение числа электронов за время dt в объеме dx вследствие разности этих потоков будет

$$I_n(x, t) dt - I_n(x + dx, t) dt = - \frac{\partial I_n}{\partial x} dx dt. \quad (8.4)$$

Полное изменение концентрации электронов в объеме dx за время dt составит величину

$$\frac{\partial n}{\partial t} dx dt = G dx dt - \frac{\partial I_n}{\partial x} dx dt - \frac{n - n_0}{\tau_n} dx dt, \quad (8.5)$$

откуда имеем:

$$\frac{\partial n}{\partial t} = G - \frac{\partial I_n}{\partial x} - \frac{n - n_0}{\tau_n}. \quad (8.6)$$

Уравнение (8.6) называют уравнением непрерывности для электронов.

Аналогично уравнение непрерывности для дырок запишется в виде

$$\frac{\partial p}{\partial t} = G - \frac{\partial I_p}{\partial x} - \frac{p - p_0}{\tau_p}. \quad (8.7)$$

Потоки электронов I_n и дырок I_p можно выразить через плотности тока, которые с учетом знака носителя равны:

$$\left. \begin{aligned} J_n &= -eI_n; \\ J_p &= eI_p. \end{aligned} \right\} \quad (8.8)$$

Когда концентрации носителей заряда являются функциями координат (x, y, z), уравнения непрерывности запишутся через соответствующие плотности токов следующим образом:

$$\frac{\partial n}{\partial t} = G + \frac{1}{e} \operatorname{div} \mathbf{J}_n - \frac{n - n_0}{\tau_n}; \quad (8.9)$$

$$\frac{\partial p}{\partial t} = G - \frac{1}{e} \operatorname{div} \mathbf{J}_p - \frac{p - p_0}{\tau_p}, \quad (8.10)$$

где

$$\operatorname{div} \mathbf{J} = \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z}.$$

Пусть в точке $x = 0$ имеется источник неравновесных носителей заряда. Рассмотрим стационарный случай. Поскольку концентрации электронов n и дырок p не изменяются во времени, т. е. $\partial n / \partial t = \partial p / \partial t = 0$, то уравнения непрерывности в одномерном случае принимают вид:

$$-\frac{1}{e} \frac{\partial J_n}{\partial x} = G - \frac{n - n_0}{\tau_n}; \quad (8.11)$$

$$\frac{1}{e} \frac{\partial J_p}{\partial t} = G - \frac{p - p_0}{\tau_p}. \quad (8.12)$$

Уравнения (8.11) и (8.12) выражают собой условия сохранения количества носителей заряда. Следовательно, в стационарном состоянии поток носителей заряда, вытекающих из объема, равен количеству носителей заряда, созданных внешним возбуждением, за вычетом носителей заряда, прорекомбинировавших в этом объеме.

8.2. ДИФФУЗИОННЫЙ И ДРЕЙФОВЫЙ ТОКИ

Рассмотрим теперь неоднородный полупроводник, в котором концентрации электронов $n(\mathbf{r})$ и дырок $p(\mathbf{r})$ изменяются от точки к точке. По этой причине в неоднородном полупроводнике должен возникнуть диффузионный ток электронов и дырок, который будет определяться диффузией носителей заряда из областей, где их концентрация больше, в области с меньшей концентрацией.

Предположим, что в полупроводнике концентрация носителей заряда возрастает в направлении оси x , как это изображено на рис. 8.2. Проведем через точку x плоскость, перпендикулярную оси x , и рассмотрим движение носителей заряда в слоях 1 и 2 толщиной dx , расположенных справа и слева от этой плоскости. В результате хаотического движения носители заряда уйдут из слоя 1, но поскольку каждый электрон может с равной вероятностью двигаться вправо и влево, половина их уйдет из слоя 1 в слой 2. Однако за это время в слой 1 придет носители заряда из слоя 2. Так как их количество в слое 2 больше, чем в 1, то обратный поток электронов будет больше прямого. Если $n(x - \frac{dx}{2})$ — средняя концентрация электронов в слое 1, а $n(x + \frac{dx}{2})$ — в слое 2, то разность концентраций электронов в этих слоях будет равна:

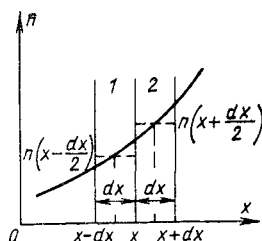


Рис. 8.2. К расчету градиента концентрации носителей заряда

$$n(x - \frac{dx}{2}) - n(x + \frac{dx}{2}) = -\frac{dn}{dx} dx. \quad (8.13)$$

Согласно (8.13) разность концентраций электронов пропорциональна градиенту их концентрации, поэтому и поток электронов I_n , возникающий в результате их диффузии в направлении x , будет пропорционален градиенту концентрации электронов в этом направлении. Его можно записать

$$I_n = -D_n \frac{dn}{dx}, \quad (8.14)$$

где D_n — коэффициент диффузии электронов.

Аналогично диффузионный поток дырок

$$I_p = -D_p \frac{dp}{dx}, \quad (8.15)$$

где D_p — коэффициент диффузии дырок.

Потоки электронов и дырок, как следует из уравнений (8.14) и (8.15), текут в сторону меньших концентраций носителей заряда.

Диффузионным потоком носителей заряда соответствуют диффузионные токи электронов $J_{n \text{ диф}}$ и дырок $J_{p \text{ диф}}$:

$$J_{n \text{ диф}} = eD_n \frac{dn}{dx}; \quad (8.16)$$

$$J_{p \text{ диф}} = -eD_p \frac{dp}{dx}. \quad (8.17)$$

В том случае, если n и p являются функциями координат (x, y, z), диффузионный ток в векторной форме имеет вид для электронов

$$J_{n \text{ диф}} = eD_n \text{grad } n(\mathbf{r}) \quad (8.18)$$

и для дырок

$$J_{p \text{ диф}} = -eD_p \text{grad } p(\mathbf{r}). \quad (8.19)$$

Диффузионный ток, возникший из-за наличия градиента концентрации носителей заряда, приведет к пространственному разделению зарядов, что вызовет появление статического электрического поля, которое создаст дрейфовые токи электронов и дырок. При термодинамическом равновесии в каждой точке полупроводника дрейфовый ток будет уравнивать диффузионный ток, поэтому суммарный ток будет равен нулю.

Допустим, что неоднородный полупроводник находится во внешнем постоянном электрическом поле напряженностью $\mathcal{E}$. Под действием этого поля электроны и дырки приобретут направленное движение, в результате чего появятся электронные и дырочные токи проводимости. Если внешнее электрическое поле слабое и не изменяет характера движения носителей заряда, то дрейфовые составляющие плотности тока запишутся на основании закона Ома в виде

$$\left. \begin{aligned} J_{n \text{ др}} &= en\mu_n \mathcal{E}; \\ J_{p \text{ др}} &= ep\mu_p \mathcal{E}. \end{aligned} \right\} \quad (8.20)$$

Полный ток будет складываться из диффузионного и дрейфового токов. Для электронов и дырок он будет равен:

$$J_n = J_{n \text{ др}} + J_{n \text{ диф}} = en\mu_n \mathcal{E} + eD_n \frac{dn}{dx}; \quad (8.21)$$

$$J_p = J_{p \text{ др}} + J_{p \text{ диф}} = ep\mu_p \mathcal{E} - eD_p \frac{dp}{dx}. \quad (8.22)$$

Таким образом, плотность общего тока J в любой точке не однородного полупроводника в любой момент времени будет определяться уравнением

$$J = J_n + J_p = e(n\mu_n + p\mu_p) \mathcal{E} + e \left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx} \right). \quad (8.23)$$

Необходимо отметить, что диффузионный ток существует только в полупроводниках. Это происходит потому, что в полупроводниках концентрации электронов и дырок могут изменяться в широких пределах при постоянной суммарной концентрации зарядов. В металлах концентрация электронов практически постоянна.

8.3. СООТНОШЕНИЕ ЭЙНШТЕЙНА

В неоднородном полупроводнике при термодинамическом равновесии ток равен нулю, т. е. $J = J_n + J_p = 0$. В этом случае токи проводимости уравнивают диффузионные токи и на основании (8.21) для электронов можно записать:

$$n\mu_n \mathcal{E}_{\text{ст}} = -D_n \frac{dn}{dx}. \quad (8.24)$$

Поскольку в полупроводнике имеется статическое электрическое поле $\mathcal{E}_{\text{ст}}$, то электроны, находящиеся в этом поле, будут обладать потенциальной энергией $U = -e\varphi$. Поэтому при отсутствии вырож-

дения концентрация электронов в зоне проводимости будет удовлетворять соотношению Больцмана вида

$$n = N_c e^{-(E_c + U - F)/kT} = n_0 e^{e\varphi/kT}, \quad (8.25)$$

где $n_0 = N_c e^{-(E_c - F)/kT}$ — равновесная концентрация электронов; φ — электростатический потенциал.

Учитывая, что $\mathcal{E}_{ст} = -\frac{d\varphi}{dx}$, и подставляя значения n и $\frac{dn}{dx}$ в уравнение (8.24), получаем:

$$-\mu_n n_0 e^{e\varphi/kT} \frac{d\varphi}{dx} = -D_n \frac{en_0}{kT} e^{e\varphi/kT} \frac{d\varphi}{dx}, \quad (8.26)$$

откуда для электронов будем иметь:

$$\frac{\mu_n}{D_n} = \frac{e}{kT}. \quad (8.27)$$

Аналогично для дырок

$$\frac{\mu_p}{D_p} = \frac{e}{kT}. \quad (8.28)$$

Уравнение, связывающее коэффициент диффузии носителей заряда, подчиняющихся статистике Максвелла, с их дрейфовой подвижностью в условиях термодинамического равновесия носит название соотношения Эйнштейна.

Как показал эксперимент, соотношение Эйнштейна применимо и к неравновесным носителям заряда. Это вполне закономерно, так как неравновесные носители заряда за малое время, намного меньше их времени жизни, обмениваясь энергией с решеткой, приходят в тепловое равновесие с решеткой, и их распределение по энергиям при отсутствии вырождения не отличается от распределения равновесных носителей заряда.

8.4. ДИФФУЗИЯ И ДРЕЙФ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА В СЛУЧАЕ МОНОПОЛЯРНОЙ ПРОВОДИМОСТИ

Рассмотрим диффузию и дрейф неравновесных носителей заряда в случае монополярной проводимости, когда свободные носители заряда возникают только в результате возбуждения их с уровней примеси. Предположим, что часть достаточно длинного однородного полупроводника, например n -типа, освещается слабопоглощающим светом. В освещенной части полупроводника при $x < 0$ (рис. 8.3, а) имеет место однородная генерация электронов в результате перевода их светом с донорных уровней в зону

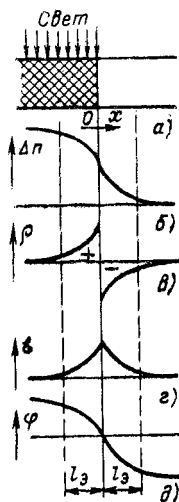


Рис. 8.3. Распределение концентрации (б), объемного заряда (в), электрического поля (е), потенциала (д) при частичном освещении донорного полупроводника (а)

проводимости. В этом случае концентрация неравновесных электронов n определяется концентрацией равновесных n_0 и избыточных электронов Δn , возбужденных с уровней донорной примеси. И если нет центров захвата, то число избыточных электронов Δn равно количеству положительных ионов донорной примеси ΔN_d^+ , т. е.

$$\Delta n = \Delta N_d^+. \quad (8.29)$$

Это равенство выражает собой условие электронейтральности в случае монополярной проводимости.

Поскольку концентрация электронов в освещенной части образца больше, чем в неосвещенной, то неравновесные электроны из освещенной части образца будут диффундировать в неосвещенную (рис. 8.3, б). Вследствие этого нарушится электронейтральность в некоторой области полупроводника и возникнет объемный заряд, а следовательно, и электрическое поле. В неосвещенной части образца, куда в результате диффузии пришли избыточные электроны, объемный заряд ρ будет отрицательным, а в освещенной области, откуда они ушли, — положительным, обусловленным ионами донорной примеси (рис. 8.3, в). Эти заряды создадут статическое электрическое поле напряженностью $\mathcal{E}_{ст}$, направленное так, что оно будет препятствовать диффузии неравновесных электронов (рис. 8.3, г). Таким образом, возникновение диффузионного тока $J_{диф} = eD_n dn/dx$ приводит к появлению статического электрического поля напряженностью $\mathcal{E}_{ст}$, а следовательно, и тока проводимости $J_{пр} = en\mu_n \mathcal{E}_{ст}$, направленного против тока диффузии. В стационарном состоянии плотность полного тока равна нулю:

$$J = J_{др} + J_{диф} = en\mu_n \mathcal{E}_{ст} + eD_n \frac{dn}{dx} = 0. \quad (8.30)$$

Из (8.30) можно определить напряженность статического электрического поля. Расчет $\mathcal{E}_{ст}$ проведем для случая малого уровня оптического возбуждения, когда концентрация избыточных электронов мала по сравнению с равновесной, т. е. $\Delta n \ll n_0$. Используя соотношение Эйнштейна, будем иметь:

$$\mathcal{E}_{ст} = - \frac{D_n}{\mu_n n} \frac{d \Delta n}{dx} = - \frac{kT}{e(n_0 + \Delta n)} \frac{d \Delta n}{dx} \approx - \frac{kT}{en_0} \frac{d \Delta n}{dx}. \quad (8.31)$$

Продифференцировав (8.31), получим:

$$\frac{d \mathcal{E}_{ст}}{dx} = - \frac{kT}{en_0} \frac{d^2 \Delta n}{dx^2}. \quad (8.32)$$

Величину $\frac{d \mathcal{E}_{ст}}{dx}$ можно найти, воспользовавшись уравнением Пуассона:

$$\frac{d \mathcal{E}_{ст}}{dx} = \frac{1}{\epsilon_r \epsilon_0} \rho = - \frac{e}{\epsilon_r \epsilon_0} \Delta n, \quad (8.33)$$

где $\rho = -e \Delta n$ — объемный заряд в неосвещенной части образца.

Из равенств (8.32) и (8.33) следует, что

$$\frac{d^2 \Delta n}{dx^2} - \frac{e^2 n_0}{\epsilon_r \epsilon_0 kT} \Delta n = 0. \quad (8.34)$$

Если ввести величину

$$l_3 = \sqrt{\varepsilon_r \varepsilon_0 k T / e^2 n_0}, \quad (8.35)$$

то (8.34) запишется следующим образом:

$$\frac{d^2 \Delta n}{dx^2} - \frac{\Delta n}{l_3^2} = 0. \quad (8.36)$$

Общее решение этого уравнения имеет вид:

$$\Delta n = C_1 e^{\alpha_1 x} + C_2 e^{\alpha_2 x}, \quad (8.37)$$

где C_1 и C_2 — постоянные, определяемые из граничных условий; α_1 , α_2 — корни характеристического уравнения, равные:

$$\alpha_1 = \frac{1}{l_3}; \quad \alpha_2 = -\frac{1}{l_3} \quad (8.38)$$

Для неосвещенной области полупроводника, в которой концентрация избыточных электронов уменьшается по мере удаления от освещенной части образца, имеет смысл только член решения с отрицательным показателем степени, поэтому

$$\Delta n = \Delta n(0) e^{-x/l_3}. \quad (8.39)$$

Таким образом, в случае монополярной проводимости концентрация избыточных носителей заряда в неосвещенной части образца по мере удаления от освещенной области уменьшается по экспоненциальному закону с постоянной спада l_3 , называемой радиусом (длиной) экранирования или дебаевским радиусом. Длина экранирования, как следует из (8.35), зависит от концентрации основных носителей заряда, поэтому ее значение может изменяться в широких пределах в зависимости от удельной проводимости полупроводника. Например, для невырожденных полупроводников, таких как германий и кремний, радиус экранирования в зависимости от степени легирования может составлять 10^{-4} — 10^{-6} см.

Воспользуемся формулой (7.29), из которой следует, что объемный заряд, введенный в полупроводник, после выключения возбуждающего света в результате тока проводимости существует в среднем в течение времени τ_μ , т. е.

$$\rho = \rho_0 e^{-t/\tau_\mu}. \quad (8.40)$$

Если плотность объемного заряда поделить на единичный заряд, то получим изменение концентрации избыточных носителей заряда во времени:

$$\Delta n = \Delta n(0) e^{-t/\tau_\mu}. \quad (8.41)$$

Из сравнения (8.39) и (8.41) следует, что распространение носителей заряда в монополярном случае на расстояние длины экранирования l_3 осуществляется в течение максвелловского времени релаксации τ_μ , которое в данном случае является эффективным временем установления диффузионно-дрейфового равновесия.

8.5. ДИФФУЗИЯ И ДРЕЙФ НЕОСНОВНЫХ ИЗБЫТОЧНЫХ НОСИТЕЛЕЙ ЗАРЯДА В ПРИМЕСНОМ ПОЛУПРОВОДНИКЕ

Рассмотрим диффузию и дрейф неосновных избыточных носителей заряда в примесном полупроводнике при наличии внешнего однородного электрического поля напряженностью $\mathcal{E}$.

Допустим, что имеется полупроводник n -типа ($p_0 \ll n_0$), поперечные размеры которого значительно меньше его длины. Очень узкая область полупроводника ($-l \leq x \leq 0$) освещается светом, так что во всем объеме освещенной области образца происходит равномерная генерация электронов и дырок (рис. 8.4). Пусть концентрация основных носителей заряда n_0 велика по сравнению с избыточной концентрацией неосновных носителей заряда Δp , время жизни которых τ_p .

В освещенной области электронного полупроводника, где возбуждаются в равном количестве избыточные электроны и дырки $\Delta n = \Delta p$, концентрация неравновесных дырок $p = p_0 + \Delta p$ будет значительно больше, чем их концентрация p_0 в неосвещенной части образца. Благодаря наличию градиента концентрации дырок возникает их диффузия. В неосвещенной части образца, куда продиффундируют дырки, появится объемный положительный заряд и электронейтральность образца нарушится. Этот объемный заряд создает статическое электрическое поле, которое вызовет перераспределение носителей заряда, так что в эту область из освещенной части образца в течение максвелловского времени релаксации будут подтянуты электроны и заряд избыточных дырок будет скомпенсирован. Полная компенсация объемного заряда будет в том случае, если концентрация избыточных электронов будет равна концентрации избыточных дырок в любой точке образца. Таким образом, *избыточные неосновные носители заряда—дырки, диффундируя в глубь электронного полупроводника, увлекают за собой равное количество основных носителей заряда—электронов, так что объемный заряд не создается, а их диффузия происходит как диффузия квазинейтрального облака неравновесных носителей, заряд в котором $\Delta n = \Delta p$. По мере продвижения в глубь полупроводника избыточные дырки и электроны будут рекомбинировать и их концентрации с расстоянием будут убывать.*

Если к такому полупроводнику приложить однородное электрическое поле $\mathcal{E}$, то при ускоряющем поле дырки будут затягивать за собой и электроны в равном количестве, а если поле будет замедлять

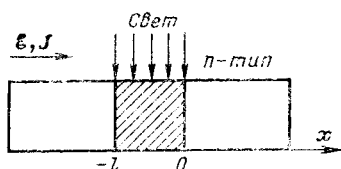


Рис. 8.4. Одномерная модель полупроводникового образца

движение дырок, то соответствующим образом произойдет и перераспределение основных носителей заряда. Следовательно, и во внешнем электрическом поле в n -материале квазинейтральное облако неравновесных носителей заряда движется в направлении движения избыточных дырок со скоростью дрейфа неосновных носителей.

Концентрация избыточных дырок Δp может быть определена в результате решения уравнения непрерывности

$$\frac{\partial p}{\partial t} = G - \frac{1}{e} \frac{\partial J_p}{\partial x} - \frac{p - p_0}{\tau_p}. \quad (8.42)$$

Проведем решение этого уравнения для стационарного случая, когда $\partial p / \partial t = 0$ в неосвещенной части полупроводника ($G = 0$). Для этого, продифференцировав выражение для плотности дырочного тока (8.22) и подставив его в (8.42), получим:

$$D_p \frac{d^2 \Delta p}{dx^2} - \mu_p \mathcal{E} \frac{d \Delta p}{dx} - \frac{\Delta p}{\tau_p} = 0. \quad (8.43)$$

Это равенство справедливо для такого внешнего электрического поля, когда D_p , μ_p и τ_p не зависят от поля.

Разделив обе части (8.43) на D_p и введя обозначения

$$L_p = \sqrt{D_p \tau_p} \quad (8.44)$$

и

$$L_{\mathcal{E}} = \tau_p \mu_p \mathcal{E}, \quad (8.45)$$

преобразуем (8.43) к виду

$$\frac{d^2 \Delta p}{dx^2} - \frac{L_{\mathcal{E}}}{L_p^2} \frac{d \Delta p}{dx} - \frac{\Delta p}{L_p^2} = 0. \quad (8.46)$$

Общим решением этого уравнения будет:

$$\Delta p = C_1 e^{\alpha_1 x} + C_2 e^{\alpha_2 x}, \quad (8.47)$$

где C_1 и C_2 — постоянные, определяемые из граничных условий; $\Delta p|_{x=0} = \Delta p|_{x=-l} = \Delta p(0)$, а при $x \rightarrow \pm \infty$ $\Delta p \rightarrow 0$; α_1 , α_2 — корни характеристического уравнения:

$$\alpha^2 - \frac{L_{\mathcal{E}}}{L_p^2} \alpha - \frac{1}{L_p^2} = 0; \quad (8.48)$$

$$\alpha_{1,2} = \frac{L_{\mathcal{E}} \pm \sqrt{L_{\mathcal{E}}^2 + 4L_p^2}}{2L_p^2}. \quad (8.49)$$

Учитывая уменьшение концентрации неосновных неравновесных носителей заряда по мере удаления от освещенной области образца, окончательно получаем:

при $x \geq 0$

$$\Delta p = C_1 e^{-x/L_1} = \Delta p(0) e^{-x/L_1}, \quad (8.50)$$

где

$$L_1 = \frac{2L_p^2}{\sqrt{L_{\mathcal{E}}^2 + 4L_p^2} - L_{\mathcal{E}}}, \quad (8.51)$$

а при $x \leq -l$

$$\Delta p = C_2 e^{x/L_2} = \Delta p(0) e^{x/L_2}, \quad (8.52)$$

где

$$L_2 = \frac{2L_p^2}{\sqrt{L_{\mathcal{E}}^2 + 4L_p^2} + L_{\mathcal{E}}}. \quad (8.53)$$

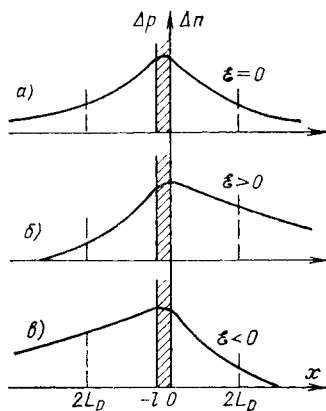


Рис. 8.5. Распределение концентрации избыточных носителей заряда вдоль образца

Таким образом, по обе стороны от освещенной области образца концентрация избыточных неосновных носителей заряда снижается по экспоненциальному закону с постоянными спада L_1 и L_2 , которые называют длиной затягивания.

Проведем анализ полученных выражений. Пусть внешнее электрическое поле отсутствует ($\mathcal{E} = 0$). В отсутствие внешнего электрического поля имеет место только диффузия избыточных дырок. Так как $L_{\mathcal{E}} = \tau_p \mu_p \mathcal{E} = 0$, то избыточная концентрация неосновных носителей заряда — дырок в результате рекомбинации будет изменяться с расстоянием по закону

$$\Delta p = \Delta p(0) e^{-x/L_p}. \quad (8.54)$$

Из этого выражения следует, что экспоненциальный спад концентрации избыточных неосновных носителей заряда, обусловленный рекомбинацией, будет симметричным по обе стороны от освещенной части образца (рис. 8.5, а) и определяется величиной L_p , называемой диффузионной длиной неосновных носителей заряда — в данном случае дырок. L_p — это то среднее расстояние, на которое смещаются неравновесные дырки при диффузии за время жизни. Следовательно, в отсутствие внешнего электрического поля избыточная концентрация дырок в результате рекомбинации будет изменяться с расстоянием по экспоненциальному закону (8.54). Это распределение справедливо и для неравновесных электронов, так как при решении задачи предполагалось выполнение электронейтральности, по которому $\Delta n = \Delta p$.

Поскольку имеет место только диффузия избыточных носителей заряда, то ток является диффузионным током, для которого будем иметь:

$$J_{p \text{ диф}} = -eD_p \frac{dp}{dx} = \frac{eD_p}{L_p} \Delta p = \frac{eL_p}{\tau_p} \Delta p = J_{p \text{ диф}}(0) e^{-x/L_p}, \quad (8.55)$$

т. е. диффузионный ток избыточных носителей заряда изменяется с координатой по тому же закону, по которому изменяется избыточная концентрация дырок.

Величину v_D , равную:

$$v_D = L_p / \tau_p = D_p / L_p, \quad (8.56)$$

называют диффузионной скоростью. Численно она равна скорости, с которой неравновесные дырки за время жизни проходят путь, равный диффузионной длине.

Рассмотрим теперь случай, когда $\mathcal{E} \neq 0$. Если $L_{\mathcal{E}} \ll 2L_p$, или $eL_p \mathcal{E} \ll 2kT$ (6.183), то $L_1 = L_2 = L_p$. Следовательно, в слабом

электрическом поле, когда $\mathcal{E} \ll \mathcal{E}_c$, где $\mathcal{E}_c = 2kT/eL_p$ — критическое поле, распределение избыточной концентрации неосновных носителей заряда определяется диффузией и описывается уравнением (8.54).

При наличии сильного внешнего электрического поля, когда $|L_{\mathcal{E}}| > 2L_p$, постоянные спада L_1 и L_2 отличаются от диффузионной длины L_p . В зависимости от направления электрического поля ($\mathcal{E} > 0$ и $\mathcal{E} < 0$) они будут больше или меньше L_p . Например, при $\mathcal{E} > 0$ в неосвещенной части для областей $x < -l$ и $x > 0$

$$L_2 = \frac{2L_p^2}{\sqrt{L_{\mathcal{E}}^2 + 4L_p^2} + L_{\mathcal{E}}} < L_p < \frac{2L_p^2}{\sqrt{L_{\mathcal{E}}^2 + 4L_p^2} - L_{\mathcal{E}}} = L_1. \quad (8.57)$$

Отсюда следует, что внешнее электрическое поле искажает симметрию в распределении избыточной концентрации носителей заряда.

Рассмотрим теперь случай таких больших полей $\mathcal{E} \gg \mathcal{E}_c$, для которых выполняется неравенство $L_{\mathcal{E}}^2 \gg 4L_p^2$. Это условие согласно (8.45) и (8.56) можно записать следующим образом:

$$L_{\mathcal{E}}^2/4L_p^2 = (\tau_p \mu_p \mathcal{E})^2/4L_p^2 = \tau_p^2 v^2/4L_p^2 = v^2/4v_D^2 \gg 1, \quad (8.58)$$

где $v = \mu_p \mathcal{E}$ — скорость дрейфа.

Величина $L_{\mathcal{E}}$, численно равная пути, проходимому неравновесным носителем заряда во время жизни со скоростью дрейфа, называется длиной дрейфа. Следовательно, напряженность поля $\mathcal{E}$ будет большой, если длина дрейфа намного превосходит диффузионную длину.

Определим длины затягивания L_1 и L_2 в случае, когда $\mathcal{E} > 0$. Для области $x > 0$ из выражения (8.51) можно записать:

$$L_1 = \frac{2L_p^2}{L_{\mathcal{E}}} \frac{1}{\sqrt{1 + 4L_p^2/L_{\mathcal{E}}^2} - 1} \approx \frac{2L_p^2}{L_{\mathcal{E}}} \frac{1}{2L_p^2/L_{\mathcal{E}}^2} = L_{\mathcal{E}}, \quad (8.59)$$

т. е. длина затягивания равна длине дрейфа

$$L_1 = L_{\mathcal{E}} = \tau_p \mu_p \mathcal{E} = \tau_p v \quad (8.60)$$

и распределение избыточной концентрации дырок определяется выражением

$$\Delta p_1 = \Delta p(0) e^{-x/L_1} = \Delta p(0) e^{-x/\tau_p \mu_p \mathcal{E}}, \quad (8.61)$$

т. е. так же, как и в случае только диффузии, избыточная концентрация дырок спадает с ростом x экспоненциально, но только теперь с постоянной спада $L_1 = L_{\mathcal{E}}$. Поскольку постоянная спада $L_1 > L_p$, то L_1 называется также «диффузионной длиной по полю».

Из сравнений (8.54) и (8.61) следует, что при $L_{\mathcal{E}}^2 \gg 4L_p^2$ в области полупроводника при $x > 0$ избыточная концентрация носителей заряда больше (правая часть рис. 8.5, б), чем в отсутствие внешнего электрического поля. Следовательно, при прохождении тока, вызванного сильным электрическим полем, т. е. когда скорость дрейфа намного больше диффузионной скорости ($v \gg v_D$), при $\mathcal{E} > 0$ избыточные дырки в электронном полупроводнике затягиваются полем в область полупроводника $x > 0$ и полупроводник обогащается

неосновными носителями заряда в большем количестве, чем при наличии только диффузии в отсутствие внешнего электрического поля ($\mathcal{E} = 0$). Это явление носит название **и н ж е к ц и и** неосновных носителей заряда.

Для дырочного полупроводника инжекция электронов будет наблюдаться при $\mathcal{E} < 0$.

Рассмотрим теперь распределение избыточных дырок для области $x < -l$ в случае $\mathcal{E} > 0$. Теперь длина затягивания

$$L_2 = \frac{2L_p^2}{L_{\mathcal{E}}} \frac{1}{\sqrt{1 + 4L_p^2/L_{\mathcal{E}}^2} + 1} \approx \frac{2L_p^2}{L_{\mathcal{E}}} \frac{1}{2 + 2L_p^2/L_{\mathcal{E}}^2} \approx \frac{L_p^2}{L_{\mathcal{E}}} \quad (8.62)$$

и на основании (8.52) получим:

$$\Delta p_2 = \Delta p(0) e^{-x/L_2} = \Delta p(0) e^{-L_{\mathcal{E}}/L_p^2}. \quad (8.63)$$

Очевидно, что с ростом напряженности поля $L_{\mathcal{E}}$ увеличивается, а L_2 уменьшается. Поскольку $L_2 < L_p$, то L_2 называют «д и ф ф у з и о н н о й д л и н о й п р о т и в п о л я» и для области $x < -l$ во всех точках полупроводника, например, на расстоянии $2L_p$ $\Delta p_2 < \Delta p$. Это значит, что при $\mathcal{E} > 0$ с ростом напряженности электрического поля объем электронного полупроводника для области $x < -l$ обедняется неосновными носителями заряда (левая часть рис. 8.5, б). Это явление носит название **э к с к л ю з и и** неосновных носителей заряда.

Для дырочного полупроводника эксклюзия электронов будет иметь место при $\mathcal{E} < 0$.

При изменении направления внешнего электрического поля $\mathcal{E} < 0$ в области полупроводника $x > 0$ будет иметь место уменьшение концентрации избыточных носителей заряда, а в области $x < -l$ — их увеличение (рис. 8.5, в). Эти явления соответственно называются **э к с т р а к ц и е й** и **а к к у м у л я ц и е й** неравновесных носителей заряда. В дырочном полупроводнике они будут наблюдаться при $\mathcal{E} > 0$.

8.6. ДИФФУЗИЯ И ДРЕЙФ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКЕ С ПРОВОДИМОСТЬЮ, БЛИЗКОЙ К СОБСТВЕННОЙ

Рассмотрим теперь диффузию и дрейф неравновесных носителей заряда в полупроводнике, в котором имеется сравнимое количество равновесных электронов и дырок, т. е. в полупроводнике с проводимостью, близкой к собственной. Будем считать, что неравновесные носители заряда возбуждаются в результате переходов «зона—зона». И если нет захвата ловушками электронов и дырок, то $\Delta n = \Delta p$.

Допустим, что в таком однородном полупроводнике существует неоднородное вдоль оси x распределение неравновесных носителей заряда $n(x) = n_0 + \Delta n(x)$ и $p(x) = p_0 + \Delta p(x)$, вызванное, например, локальным освещением образца. Будем считать, что уровень возбуждения низкий, так что $\Delta n = \Delta p \ll (n_0 + p_0)$. Благодаря на-

личию внешнего электрического поля и градиента концентраций носителей заряда возникнут диффузионные и дрейфовые токи электронов и дырок, которые будут описываться (8.21) и (8.22), и плотность тока будет равна

$$J = J_n + J_p = e(n\mu_n + p\mu_p)\mathcal{E} + e\left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx}\right). \quad (8.64)$$

В изолированном полупроводнике в стационарном состоянии полный ток равен нулю ($J = 0$). При этом в каждой точке образца диффузионные токи уравниваются дрейфовыми токами, обусловленными статическим электрическим полем напряженностью $\mathcal{E}_{ст}$, которую найдем из (8.64), если положить $J = 0$:

$$\mathcal{E}_{ст} = \frac{D_p \frac{dp}{dx} - D_n \frac{dn}{dx}}{n\mu_n + p\mu_p} = \frac{D_p \frac{dp}{dx} - D_n \frac{dn}{dx}}{(n_0 + \Delta n)\mu_n + (p_0 + \Delta p)\mu_p}. \quad (8.65)$$

Из (8.65) следует, что при отсутствии электрического тока ($J = 0$) возникающее в результате диффузии носителей заряда статическое электрическое поле $\mathcal{E}_{ст}$ будет тем меньше, чем больше равновесные концентрации электронов n_0 и дырок p_0 . Если удельная электрическая проводимость полупроводника достаточно велика и концентрации избыточных носителей заряда Δn и Δp значительно меньше концентрации равновесных носителей заряда, т. е. $\Delta n \ll n_0$ и $\Delta p \ll p_0$, то в первом приближении можно считать, что статическое электрическое поле отсутствует и имеет место условие электронейтральности, которое обеспечивается тем, что в каждой точке полупроводника $\Delta n = \Delta p$. Это значит, что диффундирующие носители заряда увлекают с собой в процессе диффузии и носители заряда противоположного знака в равном количестве. В этом случае процесс диффузии избыточных электронов и дырок происходит свободно, как диффузия избыточных нейтральных пар электрон—дырка, характеризующихся одним временем жизни τ .

Предположим, что образец, в котором созданы неравновесные электронно-дырочные пары, находится во внешнем однородном электрическом поле $\mathcal{E}$, значительно превосходящем внутреннее статическое поле, т. е. $\mathcal{E} \gg \mathcal{E}_{ст}$. Для простоты ограничимся одномерным случаем и будем считать, что градиент концентрации и внешнее электрическое поле направлены вдоль оси x . Тогда уравнения непрерывности и уравнение для плотности токов должны быть записаны как для электронов, так и для дырок:

$$\left. \begin{aligned} \frac{\partial \Delta n}{\partial t} &= \frac{1}{e} \frac{\partial J_n}{\partial x} - \frac{\Delta n}{\tau}; \\ \frac{\partial \Delta p}{\partial t} &= -\frac{1}{e} \frac{\partial J_p}{\partial x} - \frac{\Delta p}{\tau}; \end{aligned} \right\} \quad (8.66)$$

$$\left. \begin{aligned} J_n &= \sigma_n \mathcal{E} + e D_n \frac{d\Delta n}{dx}; \\ J_p &= \sigma_p \mathcal{E} - e D_p \frac{d\Delta p}{dx}; \end{aligned} \right\} \quad (8.67)$$

где $\sigma_n = en\mu_n$ и $\sigma_p = ep\mu_p$ — электронная и дырочная составляющие удельной проводимости.

Под воздействием внешнего электрического поля пары электрон—дырка будут дрейфовать с постоянной скоростью. При этом совместная диффузия и дрейф электронов и дырок при условии электро-нейтральности образца будут характеризоваться эффективной дрейфовой подвижностью $\mu_{\mathcal{E}}$ и эффективным коэффициентом диффузии D , одинаковым для электронов и дырок. Величины $\mu_{\mathcal{E}}$ и D называют также амбиполярной дрейфовой подвижностью и коэффициентом амбиполярной диффузии.

Для того чтобы определить $\mu_{\mathcal{E}}$ и D , запишем уравнения непрерывности (8.66), подставив в них значения J_n и J_p из (8.67):

$$\left. \begin{aligned} \frac{\partial \Delta n}{\partial t} &= D_n \frac{\partial^2 \Delta n}{\partial x^2} + \mu_n \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau}; \\ \frac{\partial \Delta p}{\partial t} &= D_p \frac{\partial^2 \Delta p}{\partial x^2} - \mu_p \mathcal{E} \frac{\partial \Delta p}{\partial x} - \frac{\Delta p}{\tau}. \end{aligned} \right\} \quad (8.68)$$

Умножим (8.68) соответственно на σ_p и σ_n и сложим оба уравнения. В результате, учитывая, что $\Delta n = \Delta p$, и используя соотношения Эйнштейна, получаем:

$$\frac{\partial \Delta n}{\partial t} = \frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} \frac{\partial^2 \Delta n}{\partial x^2} + \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau}. \quad (8.69)$$

Для стационарного случая, когда $\frac{\partial \Delta n}{\partial t} = 0$, (8.69) запишется в виде

$$\frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} \frac{\partial^2 \Delta n}{\partial x^2} + \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau} = 0. \quad (8.70)$$

Уравнение (8.70) по форме совпадает с выражением (8.43) для диффузии и дрейфа неосновных избыточных носителей заряда. Разница между ними лишь в том, что вместо коэффициента диффузии при второй производной и подвижности при первой производной в (8.70) стоят сложные величины, которые соответственно могут быть обозначены через D и $\mu_{\mathcal{E}}$.

Учитывая, что при $n \approx n_0$ $p \approx p_0$, а это справедливо, когда $\Delta n \ll n_0$ и $\Delta p \ll p_0$, и используя соотношение Эйнштейна для электронов и дырок ($\mu_n/D_n = \mu_p/D_p = e/kT$), коэффициент амбиполярной диффузии можно записать в виде

$$D = \frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} = \frac{n_0 + p_0}{\frac{n_0}{D_p} + \frac{p_0}{D_n}} = \frac{kT}{e} \frac{n_0 + p_0}{\frac{n_0}{\mu_p} + \frac{p_0}{\mu_n}}, \quad (8.71)$$

а амбиполярную дрейфовую подвижность в виде

$$\mu_{\mathcal{E}} = \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} = \frac{p_0 - n_0}{\frac{n_0}{\mu_p} + \frac{p_0}{\mu_n}}. \quad (8.72)$$

Если воспользоваться соотношением Эйнштейна, то коэффициент амбиполярной диффузии D можно представить в виде

$$D = \frac{kT}{e} \mu_D. \quad (8.73)$$

Из сопоставления с равенством (8.73) следует, что в (8.71) роль подвижности играет величина, равная:

$$\mu_D = \frac{\frac{n_0}{\mu_n} + \frac{p_0}{\mu_p}}{\frac{n_0}{\mu_n} + \frac{p_0}{\mu_p}} \quad (8.74)$$

и называемая амбиполярной диффузионной подвижностью.

Из (8.72) следует, что $\mu_{\mathcal{E}}$ может изменять знак в зависимости от соотношения n_0 и p_0 . Это определяется тем, что во внешнем электрическом поле $\mu_{\mathcal{E}}$ характеризует скорость дрейфа квазинейтрального облака неравновесных носителей заряда, которое движется в направлении движения неосновных носителей. В электронном полупроводнике $\mu_{\mathcal{E}} > 0$ и облако движется в направлении движения дырок, а в p -материале облако движется в противоположном направлении — в направлении движения электронов.

Для собственного полупроводника, у которого $n_0 = p_0 = n_i$, имеем:

$$D = 2 \frac{D_n D_p}{D_n + D_p} = 2 \frac{kT}{e} \frac{\mu_n \mu_p}{\mu_n + \mu_p}; \quad (8.75)$$

$$\mu_D = 2 \frac{\mu_n \mu_p}{\mu_n + \mu_p}; \quad (8.76)$$

$$\mu_{\mathcal{E}} = 0. \quad (8.77)$$

Из (8.75) и (8.77) следует, что для собственного полупроводника величина коэффициента амбиполярной диффузии зависит только от коэффициентов диффузии (или подвижностей) электронов и дырок. Равенство нулю амбиполярной дрейфовой подвижности свидетельствует о том, что внешнее электрическое поле в собственном полупроводнике, когда $\Delta n = \Delta p$, не влияет на пространственное распределение носителей заряда.

Для примесных полупроводников, у которых концентрация основных носителей заряда значительно превышает концентрацию неосновных носителей заряда ($n_0 \gg p_0$ или $p_0 \gg n_0$), коэффициент амбиполярной диффузии D , амбиполярная диффузионная подвижность μ_D и амбиполярная дрейфовая подвижность $\mu_{\mathcal{E}}$ равны соответственно коэффициенту диффузии и подвижности неосновных носителей заряда. Например, для полупроводника n -типа ($n_0 \gg p_0$)

$$D = D_p,$$

а

$$\mu_D = |\mu_{\mathcal{E}}| = \mu_p. \quad (8.78)$$

Следовательно, в примесном полупроводнике диффузия и дрейф избыточных носителей заряда определяются соответственно коэффициентом диффузии и подвижностью неосновных носителей заряда,

КОНТАКТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

9.1. ПОЛУПРОВОДНИК ВО ВНЕШНЕМ ЭЛЕКТРИЧЕСКОМ ПОЛЕ

В основе контактных явлений лежат физические процессы, протекающие в полупроводнике под действием электрического поля, возникающего на контакте. Для понимания этих явлений необходимо знание свойств полупроводника, находящегося во внешнем электрическом поле. С этой целью проведем рассмотрение свойств электронного полупроводника, внесенного в однородное электрическое поле конденсатора (рис. 9.1).

В полупроводнике в отсутствие внешнего электрического поля объемный заряд равен нулю. При наличии внешнего электрического поля в полупроводнике произойдет перераспределение носителей заряда, в результате чего в нем появится объемный заряд, плотность которого $\rho(\mathbf{r})$, и электрическое поле напряженностью $\mathcal{E}(\mathbf{r})$. Изменение распределения концентрации носителей заряда, обуславливающее появление объемного заряда, будет иметь место в области, прилегающей к поверхности полупроводника. Объемный заряд будет экранировать внешнее электрическое поле, вследствие чего оно проникнет только в приповерхностный слой полупроводника.

При таком приложении внешнего электрического поля, как это изображено на рис. 9.1, в приповерхностной области полупроводника будет повышенная концентрация электронов (рис. 9.2, а) и, следовательно, возникнет отрицательный объемный заряд (рис. 9.2, б), который будет уменьшаться с увеличением расстояния от поверхности в глубь образца. Отрицательный объемный заряд создаст электрическое поле, напряженность которого $\mathcal{E}_s$ будет максимальной на поверхности полупроводника (рис. 9.2, в). Это электрическое поле изменит потенциальную энергию электрона (рис. 9.2, д) на величину, равную $U(\mathbf{r}) = -e\varphi(\mathbf{r})$, где $\varphi(\mathbf{r})$ — электростатический потенциал контактного поля (рис. 9.2, з). Следовательно, электрическое поле вызовет искривление зон энергии полупроводника таким образом, что

$$\left. \begin{aligned} E_c(\mathbf{r}) &= E_c - U(\mathbf{r}); \\ E_v(\mathbf{r}) &= E_v - U(\mathbf{r}). \end{aligned} \right\} \quad (9.1)$$

При этом изменится положение всех уровней энергии, в том числе и уровней примеси, лежащих в запрещенной зоне (рис. 9.2, е).

Поскольку полупроводник находится в состоянии термодинамического равновесия, то положение уровня Ферми постоянно (см. § 5.2), поэтому расстояние между уровнем Ферми и разрешенными зонами энергии изменится. Если это расстояние без поля было:

$$E_c - F \text{ и } F - E_v, \quad (9.2)$$

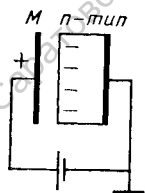


Рис. 9.1. Полупроводник в однородном электрическом поле

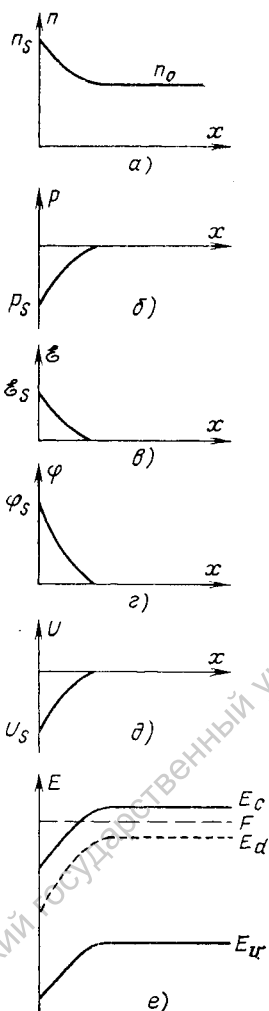


Рис. 9.2. Изменение концентрации свободных электронов (а), объемного заряда (б), напряженности поля (в), потенциала поля (г), потенциальной энергии электронов (д) и искривление зон донорного полупроводника (е) во внешнем электрическом поле

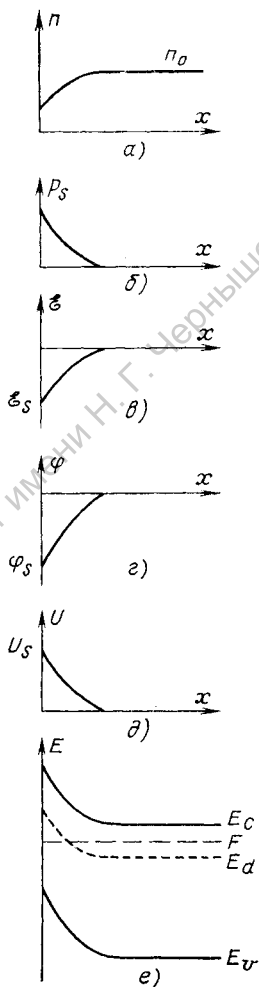


Рис. 9.3. Концентрация свободных электронов (а), плотность объемного заряда (б), напряженность электрического поля (в), потенциал поля (г), потенциальная энергия электронов (д) и искривление зон донорного полупроводника (е) при изменении направления внешнего электрического поля

то при наличии поля оно будет:

$$[E_c - U(r)] - F \text{ и } F - [E_v - U(r)]. \quad (9.3)$$

Из сравнения (9.3) и (9.2) следует, что если расстояние между E_c и F уменьшается на величину $U(r)$, то между F и E_v увеличивается на ту же величину. Изменение расстояния между F и краями зон энергии должно привести к изменению распределения электронов по уровням. В рассматриваемом случае, как это изображено на рис. 9.2, *е*, вдали от поверхности полупроводник обладает электронной электропроводимостью, невырожден и на уровнях донорной примеси имеются электроны (поскольку уровень Ферми лежит выше уровня примеси), а в приповерхностной области полупроводник, оставаясь электронным, становится вырожденным, поскольку уровень Ферми расположен теперь в зоне проводимости.

При изменении направления приложенного поля в приповерхностной области электронного полупроводника концентрация электронов будет снижена по сравнению с концентрацией в объеме полупроводника (рис. 9.3, *а*), а объемный заряд (рис. 9.3, *б*) и изменение потенциальной энергии (рис. 9.3, *д*) по знаку будут противоположные. В данном случае, как это изображено на рис. 9.3, *е*, в приповерхностной области произошло даже изменение типа основных носителей заряда — полупроводник из электронного превратился в дырочный, т. е. возник *и н в е р с н ы й*, или *о б р а щ е н н ы й*, *с л о й*. У такого полупроводника на некотором расстоянии от поверхности будет область с собственной проводимостью или *і-слой* с уровнем Ферми, лежащим в середине запрещенной зоны. *Эта область полупроводника, расположенная около і-го слоя, в которой меняется тип электропроводности, носит название физического р-п перехода.* Он исчезает при снятии внешнего электрического поля.

Рассмотрение влияния внешнего электрического поля проведем для одномерного невырожденного полупроводника *n*-типа, энергетическая схема которого представлена на рис. 9.3, *е*. Как известно, напряженность электрического поля $\mathcal{E}$ связана с объемным зарядом ρ уравнением Пуассона

$$\frac{d\mathcal{E}}{dx} = \frac{1}{\varepsilon_r \varepsilon_0} \rho(x). \quad (9.4)$$

Если напряженность поля выразить через градиент потенциала $\mathcal{E} = -d\varphi/dx$, то уравнение Пуассона будет иметь вид:

$$\frac{d^2\varphi}{dx^2} = -\frac{1}{\varepsilon_r \varepsilon_0} \rho(x). \quad (9.5)$$

Пусть концентрация электронов в объеме полупроводника ($x \rightarrow \infty$) равна n_0 . Поскольку полупроводник невырожден, то в его приповерхностной области

$$n = N_c e^{-(E_c + U - F)/kT} = n_0 e^{-U/kT}. \quad (9.6)$$

В приповерхностном слое объемный заряд определяется положительными ионами донорной примеси и свободными электронами,

Будем считать, что донорная примесь полностью ионизована, т. е. $N_d^+ = n_0$. Тогда объемный заряд в приповерхностном слое будет равен:

$$\rho = e(N_d^+ - n) = e(n_0 - n) = en_0(1 - e^{-U/kT}). \quad (9.7)$$

Ограничимся случаем малого искривления зон под действием внешнего электрического поля, т. е. когда $|U| \ll kT$. Разлагая выражение $e^{-U/kT}$ в ряд и ограничиваясь первым членом разложения, на основании соотношения (9.7) будем иметь:

$$\rho = \frac{en_0 U}{kT} = -\frac{e^2 n_0}{kT} \varphi. \quad (9.8)$$

Обозначим $l_s^2 = \epsilon_r \epsilon_0 kT / e^2 n_0$, тогда (9.5) примет вид:

$$\frac{d^2 \varphi}{dx^2} - \frac{1}{l_s^2} \varphi = 0. \quad (9.9)$$

Это уравнение имеет решение

$$\varphi = A_1 e^{-x/l_s} + A_2 e^{x/l_s}. \quad (9.10)$$

Так как при $x \rightarrow \infty$ $\varphi \rightarrow 0$, значит, $A_2 = 0$, а в точке $x = 0$ $\varphi = -\varphi_s$ и $A_1 = -\varphi_s$, поэтому в приповерхностной области потенциал

$$\varphi(x) = -\varphi_s e^{-x/l_s}, \quad (9.11)$$

напряженность поля

$$\mathcal{E}(x) = -\frac{d\varphi}{dx} = -\frac{\varphi_s}{l_s} e^{-x/l_s} = -\mathcal{E}_s e^{-x/l_s}, \quad (9.12)$$

потенциальная энергия электронов

$$U(x) = -e\varphi(x) = e\varphi_s e^{-x/l_s} = U_s e^{-x/l_s}, \quad (9.13)$$

а плотность объемного заряда на поверхности

$$\rho_s = \frac{en_0}{kT} U_s. \quad (9.14)$$

Таким образом, при внесении полупроводника в электрическое поле в приповерхностной области его происходят искривление зон энергии и изменение концентрации электронов и дырок. При $U_s > 0$ зоны искривляются кверху и увеличивается концентрация дырок (рис. 9.3, е). В этом случае у полупроводника *n*-типа приповерхностный слой обогащается неосновными носителями заряда, а у *p*-типа — основными носителями заряда. При $U_s < 0$ зоны искривляются вниз (рис. 9.2, е) и увеличивается концентрация электронов, т. е. у полупроводника *p*-типа имеет место обогащение основными носителями заряда, а у *n*-типа — обогащение неосновными носителями заряда.

В (9.10)–(9.13) l_s есть дебаевская длина экранирования. Следовательно, l_s — это то расстояние, на котором напряженность внеш-

него электрического поля, проникающего в вещество, уменьшается в $e \approx 2.7$ раза. Для металла, у которого $n_0 = 10^{22} \text{ см}^{-3}$ и $\epsilon_r = 1$, при комнатной температуре дебаевская длина экранирования составляет примерно 10^{-8} см . В германий с $n_0 = 10^{14} \text{ см}^{-3}$ и $\epsilon_r = 16$ электрическое поле проникает на расстояние 4 мкм.

9.2. ТЕРМОЭЛЕКТРОННАЯ РАБОТА ВЫХОДА

Как известно, чтобы перевести электрон из твердого тела в вакуум, необходимо затратить некоторую энергию. Следовательно, на границе твердого тела и вакуума существует энергетический барьер, препятствующий выходу электронов из кристалла. Поэтому покинуть твердое тело могут лишь те электроны, которые имеют энергию, достаточную для преодоления этого барьера. Очевидно, что чем выше температура, тем больше будет электронов, способных преодолеть потенциальный барьер. Явление выхода из вещества электронов вследствие теплового возбуждения называют термоэлектронной эмиссией.

Определим плотность тока термоэлектронной эмиссии электронов из невырожденного полупроводника. Для этого необходимо подсчитать число электронов, способных преодолеть потенциальный барьер и выйти из полупроводника. Обозначим через $E_a = E_{\text{вак}}$ энергию электрона, вышедшего из полупроводника в вакуум и покоящегося относительно образца. Тогда для перевода электрона со дна зоны проводимости полупроводника в вакуум без сообщения ему скорости потребуется энергия χ (рис. 9.4), равная:

$$\chi = E_a - E_c. \quad (9.15)$$

Энергия χ есть энергия электронного сродства. Численно она равна работе, необходимой для перевода электрона со дна зоны проводимости в вакуум без сообщения ему кинетической энергии.

Для преодоления прямоугольного потенциального барьера высотой χ кинетическая энергия электрона, имеющего скорость v_x , должна быть не меньше глубины потенциального барьера, т. е.

$$m^* v_x^2 / 2 \geq \chi. \quad (9.16)$$

Количество квантовых состояний для кристалла единичного объема в интервале скоростей от v до $v + dv$ определяется выражением (см. § 4.1)

$$dZ = 2 (m^* / h)^3 dv_x dv_y dv_z. \quad (9.17)$$

Поскольку для электронов, способных покинуть полупроводник, $E - F \gg kT$, то их число будет:

$$dn = \int_0 dZ = 2 (m^* / h)^3 \times \int e^{-(E-F)/kT} dv_x dv_y dv_z. \quad (9.18)$$

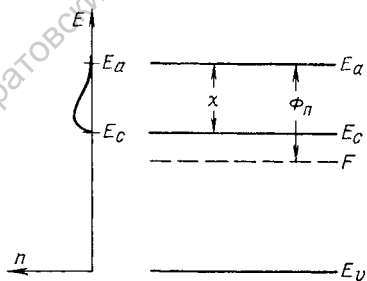


Рис. 9.4. Термоэлектронная работа выхода электронного полупроводника

Будем считать, что все электроны, преодолевшие потенциальный барьер, не возвращаются в полупроводник. Если к поверхности кристалла движется поток электронов со скоростью v_x , равный dnv_x , то плотность тока, текущего из вакуума в полупроводник, будет

$$J = e \int_{v_{x \min}}^{\infty} \int_{-\infty}^{\infty} dnv_x = 2e (m^*/h)^3 e^{F/kT} \int_{v_{x \min}}^{\infty} \int_{-\infty}^{\infty} e^{-E/kT} v_x dv_x dv_y dv_z. \quad (9.19)$$

Учитывая, что полная энергия электрона $E = E_c + \frac{m^*}{2} (v_x^2 + v_y^2 + v_z^2)$, уравнение (9.19) можно записать в виде

$$J = \frac{2em^*{}^3}{h^3} e^{-(E_c-F)/kT} \int_{v_{x \min}}^{\infty} e^{-m^* v_x^2/2kT} v_x dv_x \times \\ \times \int_{-\infty}^{\infty} e^{-m^* v_y^2/2kT} dv_y \int_{-\infty}^{\infty} e^{-m^* v_z^2/2kT} dv_z. \quad (9.20)$$

Для вычисления J используем соотношение

$$\int_{-\infty}^{\infty} e^{-\alpha y^2} dy = \sqrt{\frac{\pi}{\alpha}}, \quad (9.21)$$

а интегрирование по v_x с учетом (9.16) дает:

$$\int_{v_{x \min}}^{\infty} e^{-m^* v_x^2/2kT} v_x dv_x = \frac{kT}{m^*} e^{-m^* v_{x \min}^2/2kT} = \frac{kT}{m^*} e^{-\chi/kT}. \quad (9.22)$$

На основании (9.20)–(9.22) получим выражение для плотности тока термоэлектронной эмиссии:

$$J = \frac{4\pi em^* k^2}{h^3} T^2 e^{-(E_a-F)/kT} = AT^2 e^{-\Phi/kT}, \quad (9.23)$$

где

$$A = 4\pi em^* k^2/h^3;$$

$$\Phi = \chi + E_c - F = E_a - F. \quad (9.24)$$

Энергия Φ , получившая название термоэлектронной работы выхода, согласно (9.24) равна разности между энергией покоящегося электрона в вакууме у поверхности образца полупроводника и уровнем Ферми в данном полупроводнике. Как у металлов, так и у полупроводников Φ составляет обычно несколько электрон-вольт.

Поскольку в полупроводнике положение уровня Ферми зависит от температуры, природы и концентрации примеси, то и термоэлектронная работа выхода также будет определяться этими параметрами. Выразим работу выхода для различных полупроводников, используя (9.24) и соответствующие соотношения для уровня Ферми, которые были получены в § 4.6 и 4.7.

Для собственного полупроводника работа выхода Φ_i равна:

$$\Phi_i = \chi + \frac{1}{2} E_g + \frac{kT}{2} \ln \left(\frac{m_n^*}{m_p^*} \right)^{3/2}. \quad (9.25)$$

Из приведенного выражения следует, что работа выхода электрона из собственного полупроводника зависит от ширины запрещенной зоны, температуры и соотношения эффективных масс электронов и дырок.

Для донорного полупроводника при слабой ионизации примеси

$$\Phi_n = \chi + \frac{E_c - E_d}{2} + \frac{kT}{2} \ln \frac{gN_c}{N_d}, \quad (9.26)$$

а при сильной ионизации примеси

$$\Phi_n = \chi + kT \ln (N_c/N_d). \quad (9.27)$$

Если при сильной ионизации донорной примеси термоэлектронная работа выхода определяется концентрацией примеси и температурой полупроводника, то при слабой ионизации она, кроме того, зависит еще и от глубины залегания уровня донорной примеси в запрещенной зоне и степени его вырождения.

Для акцепторного полупроводника при слабой и сильной ионизации примеси

$$\Phi_p = \chi + E_g + \frac{E_v - E_a}{2} - \frac{kT}{2} \ln \frac{gN_v}{N_a}; \quad (9.28)$$

$$\Phi_p = \chi + E_g - kT \ln \frac{N_v}{N_a}. \quad (9.29)$$

Из сравнения формул (9.26)—(9.29) следует, что работа выхода электронов из дырочного полупроводника больше, чем из электронного полупроводника той же природы.

9.3. КОНТАКТ МЕТАЛЛ—МЕТАЛЛ.

КОНТАКТНАЯ РАЗНОСТЬ ПОТЕНЦИАЛОВ

Допустим, имеем два различных изолированных металла: M_1 и M_2 . Предположим, что у металла M_1 работа выхода Φ_1 и энергия Ферми F_1 больше, чем Φ_2 и F_2 металла M_2 (рис. 9.5, а).

Рассмотрим вначале процессы, происходящие в плоскости контакта C (рис. 9.5, б). В первый момент соприкосновения металлы

начнут обмениваться электронами. Поскольку $F_1 > F_2$, то преимущественным будет переход электронов из металла M_1 в металл M_2 , поэтому на металле M_1 появится положительный заряд, а на металле M_2 — отрицательный. Следовательно, между металлами перпендикулярно плоскости контакта C

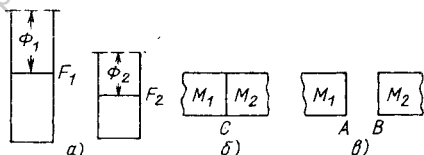
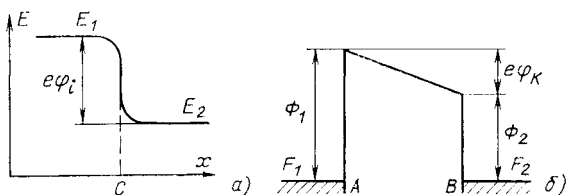


Рис. 9.5. Потенциальные барьеры двух изолированных металлов (а), контакт металлов в плоскости C (б) и на вакуумном зазоре (в)

Рис. 9.6. Внутренняя (а) и внешняя (б) контактная разность потенциалов



возникнет электрическое поле, которое будет локализовано в слое толщиной порядка 10^{-8} см. Направленный поток электронов из металла M_1 в металл M_2 прекратится в тот момент, когда уровень Ферми для всей системы будет одинаков. Разность потенциалов, установившуюся при равновесном состоянии на контакте между двумя металлами, равную ϕ_i , называют внутренней контактной разностью потенциалов.

Для определения ϕ_i необходимо подсчитать число электронов, перешедших вдоль оси x из металла M_1 в металл M_2 и наоборот. В соответствии с (9.18) из металла M_1 перейдут электроны, имеющие скорость v_{x1} , в количестве

$$dN_1 = 2 \frac{1}{m_1^* h^3} p_{x1} dp_{x1} \iint_{-\infty}^{\infty} \frac{dp_{y1} dp_{z1}}{e^{(E_1 - F_1)/kT} + 1}, \quad (9.30)$$

а количество электронов, перешедших из второго в первый металл, составит

$$dN_2 = 2 \frac{1}{m_2^* h^3} p_{x2} dp_{x2} \iint_{-\infty}^{\infty} \frac{dp_{y2} dp_{z2}}{e^{(E_2 - F_2)/kT} + 1}. \quad (9.31)$$

Так как металл M_1 на границе раздела C заряжен положительно, то его электроны приобретут дополнительную энергию, равную $-e\phi_i$ (рис. 9.6, а). Но при термодинамическом равновесии полная энергия электронов по всей системе одинакова, поэтому

$$E_1 - e\phi_i = E_2. \quad (9.32)$$

Поскольку

$$\frac{\partial E_1}{\partial p_{x1}} dp_{x1} = \frac{\partial E_2}{\partial p_{x2}} dp_{x2}, \quad (9.33)$$

то

$$\frac{1}{m_1^*} p_{x1} dp_{x1} = \frac{1}{m_2^*} p_{x2} dp_{x2}. \quad (9.34)$$

В условиях термодинамического равновесия число прямых переходов равно числу обратных переходов, т. е.

$$dN_1 = dN_2. \quad (9.35)$$

Это условие осуществимо лишь при равенстве двойных интегралов в (9.30) и (9.31). Их равенство возможно в том случае, если

$$E_1 - F_1 = E_2 - F_2 = E_1 - e\phi_i - F_2. \quad (9.36)$$

Отсюда следует, что

$$e\varphi_i = F_1 - F_2. \quad (9.37)$$

Таким образом, *внутренняя контактная разность потенциалов на контакте двух металлов в плоскости С определяется разностью энергий Ферми изолированных металлов.*

Для металлов уровень Ферми связан с концентрацией электронов выражением

$$F_0 = \frac{\hbar^2}{2m^*} \left(\frac{3n_0}{8\pi} \right)^{2/3}, \quad (9.38)$$

поэтому

$$e\varphi_i = \frac{\hbar^2}{2} \left(\frac{3}{8\pi} \right)^{2/3} \left(\frac{n_1^{2/3}}{m_1^*} - \frac{n_2^{2/3}}{m_2^*} \right). \quad (9.39)$$

Следовательно, *внутренняя контактная разность потенциалов на контакте двух металлов определяется концентрацией электронов в изолированных металлах и их эффективными массами.* Для металлов φ_i невелико и составляет 10^{-3} — 10^{-2} В.

Рассмотрим теперь процессы, происходящие между металлами в вакуумном зазоре (на участке АВ, рис. 9.5, в). Поскольку работа выхода электронов из металла M_1 больше, чем из металла M_2 , то в первый момент ток термоэлектронной эмиссии из металла M_2 будет больше, чем из металла M_1 , т. е. $J_2 > J_1$. Это приведет к тому, что на металле M_2 (в плоскости В) появится положительный заряд, а на металле M_1 (в плоскости А) — отрицательный, а следовательно, возникнет электрическое поле, препятствующее эмиссии электронов из металла M_2 . Направленный поток электронов будет существовать до тех пор, пока в системе не наступит равновесие, характеризующееся постоянством уровня Ферми (рис. 9.6, б). *Контактная разность потенциалов φ_k , возникшая между металлами в вакуумном зазоре, носит название внешней контактной разности потенциалов.*

Поскольку при термодинамическом равновесии имеет место равенство токов эмиссии

$$J_{01} = J_{02} \quad (9.40)$$

или

$$AT^2 e^{-\Phi_1/kT} = AT^2 e^{-(\Phi_2 + e\varphi_k)/kT}, \quad (9.41)$$

то

$$e\varphi_k = \Phi_1 - \Phi_2. \quad (9.42)$$

Таким образом, *внешняя контактная разность потенциалов, возникающая между металлами в вакуумном зазоре, определяется разностью термоэлектронных работ выхода этих металлов.*

9.4. КОНТАКТ МЕТАЛЛ—ПОЛУПРОВОДНИК

Допустим, что имеется контакт между металлом и невырожденным электронным полупроводником. Если уровень Ферми изолированного металла F_M лежит ниже уровня Ферми полупроводника F_n , т. е. $\Phi_M > \Phi_n$, то в первый момент соприкосновения поток электро-

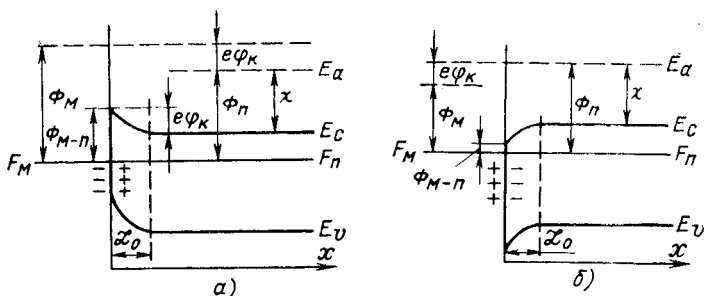


Рис. 9.7. Контакт металл—электронный полупроводник в случае $\Phi_M > \Phi_n$ (а) и $\Phi_M < \Phi_n$ (б)

нов из полупроводника превышает поток электронов из металла. В результате металл в области контакта приобретает отрицательный заряд, а полупроводник — положительный и возникшее между контактирующими образцами электрическое поле $\mathcal{E}_{ct}$ будет препятствовать переходу электронов из полупроводника в металл. Направленный поток электронов будет иметь место до тех пор, пока уровни Ферми в системе не выравняются (рис. 9.7, а) и установится равновесие, характеризующееся равенством токов термоэлектронной эмиссии

$$J_{M0} = J_{n0}. \quad (9.43)$$

Отсюда найдем, что контактная разность потенциалов, возникающая на контакте металла с полупроводником, равна

$$e\varphi_K = \Phi_M - \Phi_n. \quad (9.44)$$

Так как напряженность электрического поля в слое объемного заряда $\mathcal{E}_{ct} = \int \frac{1}{\epsilon_r \epsilon_0} \rho dx = \frac{1}{\epsilon_r \epsilon_0} Q$, то при равенстве объемных зарядов ($Q = \text{const}$) разность потенциалов в соприкасающихся телах пропорциональна толщине слоя объемного заряда ($V =$

$= - \int_0^l \mathcal{E}_{ct} dx = - \frac{Q}{\epsilon_r \epsilon_0} l$). Но толщина слоя объемного заряда в металле не превышает 10^{-8} — 10^{-7} см, а в полупроводнике может составлять 10^{-4} см и больше и поэтому разность потенциалов в слое объемного заряда металла ничтожно мала по сравнению с разностью потенциалов в полупроводнике. В силу этого можно считать, что контактная разность потенциалов φ_K практически полностью приходится на приконтактную область полупроводника. Напряженность электрического поля в приповерхностном слое, вызванного контактной разностью потенциалов, не превышает 10^6 В/см, а напряженность поля, обусловленного ионами кристаллической решетки, составляет 10^8 В/см. Поэтому небольшая добавка контакт-

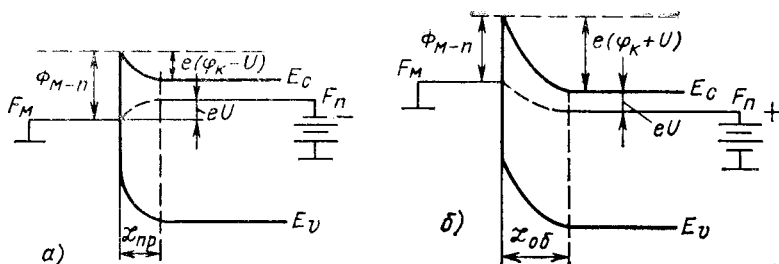


Рис. 9.8. Контакт металл-дырочный полупроводник в случае $\Phi_M > \Phi_P$ (а) и $\Phi_M < \Phi_P$ (б)

ной разности потенциалов не может изменить ширину запрещенной зоны полупроводника, но искривит его зоны энергии относительно уровня Ферми, а это вызовет изменение концентрации электронов и дырок в его приповерхностном слое.

При этом, если работа выхода электронов из полупроводника меньше работы выхода из металла, полупроводник заряжается положительно и зоны энергии в приконтактной области искривляются вверх, так что дно зоны проводимости E_C удаляется от уровня Ферми F_P , а потолок валентной зоны E_V , наоборот, приближается к уровню Ферми F_P (рис. 9.7, а и 9.8, а). Поэтому вблизи контакта концентрация электронов убывает, а количество дырок в валентной зоне возрастает по сравнению с их концентрациями в объеме полупроводника. Таким образом, в случае, когда $\Phi_M > \Phi_P$, в электронном полупроводнике концентрация дырок приконтактной области будет больше, чем в его объеме (рис. 9.7, а), т. е. возникает слой с пониженной удельной проводимостью. Приконтактный слой полупроводника с пониженной удельной проводимостью (обогащенный неосновными носителями заряда) называют заборным. У дырочного полупроводника концентрация дырок у контакта будет больше, чем в объеме образца (рис. 9.8, а), т. е. возникает слой с повышенной удельной проводимостью. Приконтактный слой полупроводника с повышенной удельной проводимостью (обогащенный основными носителями заряда) принято называть антизаборным.

Если работа выхода из полупроводника больше работы выхода из металла, т. е. $\Phi_P > \Phi_M$, полупроводник заряжается отрицательно и его зоны энергии в приконтактной области искривляются книзу, поэтому вблизи контакта число электронов в зоне проводимости увеличивается, а число дырок в валентной зоне убывает. Это значит, что в электронном полупроводнике возникнет антизаборный слой (рис. 9.7, б), а в дырочном — заборный слой (рис. 9.8, б).

При сильном обогащении приконтактной области неосновными носителями заряда, приводящего к инверсии электропроводности, возникнет физический p-n переход. Сильное обогащение приконтактного слоя основными носителями заряда может привести к вырождению полупроводника в этой области.

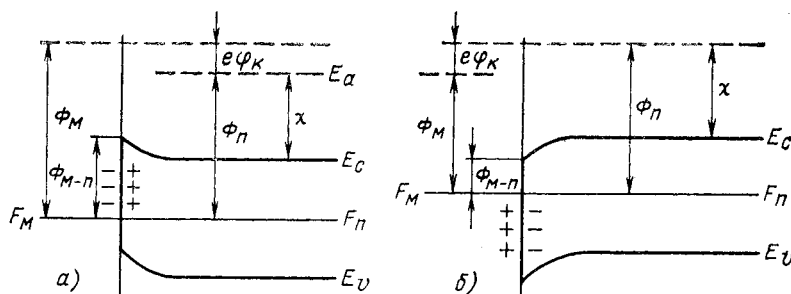


Рис. 9.9. Контакт металл—собственный полупроводник в случае $\Phi_M > \Phi_P$ (а) и $\Phi_M < \Phi_P$ (б)

У собственного полупроводника как при $\Phi_M > \Phi_P$, так и при $\Phi_M < \Phi_P$ искривление зон энергии сопровождается повышением удельной проводимости в приконтактной области (рис. 9.9).

Преобразуем (9.44) следующим образом:

$$e\varphi_K = (\Phi_M - \chi) - (\Phi_P - \chi) = \Phi_{M-P} - \Phi_{P-M}. \quad (9.45)$$

Здесь $\Phi_{M-P} = \Phi_M - \chi$ — работа, которую необходимо совершить для перевода электрона с уровня Ферми металла на дно зоны проводимости полупроводника в контакте (см. рис. 9.7—9.9); $\Phi_{P-M} = \Phi_P - \chi$ — энергия, которую необходимо затратить для перевода электрона, находящегося на дне зоны проводимости в объеме полупроводника, непосредственно в металл.

Предположим, что электрическое поле проникает в электронный полупроводник на глубину $\mathcal{L}_0$ (рис. 9.7, а). Ее определение проведем для полупроводника, у которого имеет место полная ионизация донорной примеси, т. е. $n_0 = N_D$. Поскольку в приконтактном слое энергия электронов на дне зоны проводимости равна $E_C + e\varphi(x)$, объемный заряд в этом слое согласно (9.7) определяется соотношением вида

$$\rho = en_0 [1 - e^{-e\varphi(x)/kT}]. \quad (9.46)$$

Так как контактная разность потенциалов полностью приходится на приконтактную область полупроводника, можно считать, что $|e\varphi(x)| \gg kT$. Для этого случая объемный заряд

$$\rho = en_0. \quad (9.47)$$

Это означает, что из области полупроводника протяженностью $\mathcal{L}_0$ свободные электроны вытеснены электрическим полем и в приконтактной области сосредоточен положительный заряд, обусловленный ионами донорной примеси.

Для области объемного заряда уравнение Пуассона с учетом (9.47) запишется в виде

$$\frac{d^2\varphi}{dx^2} + \frac{e}{\epsilon_r \epsilon_0} n_0 = 0. \quad (9.48)$$

Общим решением этого уравнения является функция

$$\varphi(x) = -\frac{en_0}{2\epsilon_r\epsilon_0}(\mathcal{L}_0 - x)^2 + A(\mathcal{L}_0 - x) + B. \quad (9.49)$$

Поле проникает в полупроводник только на глубину $\mathcal{L}_0$, а поэтому уравнение (9.49) должно удовлетворять граничным условиям

$$\varphi(\mathcal{L}_0) = 0;$$

$$\mathcal{E}(\mathcal{L}_0) = -\left.\frac{d\varphi}{dx}\right|_{x=\mathcal{L}_0} = 0. \quad (9.50)$$

Подставляя граничные условия в (9.49), получаем:

$$\left. \begin{aligned} \varphi(\mathcal{L}_0) = B = 0; \\ \left.\frac{d\varphi}{dx}\right|_{x=\mathcal{L}_0} = -A = 0. \end{aligned} \right\} \quad (9.51)$$

Следовательно, в приконтактной области электронного полупроводника (рис. 9.7, а) электростатический потенциал в зависимости от координаты x меняется как

$$\varphi(x) = -\frac{en_0}{2\epsilon_r\epsilon_0}(\mathcal{L}_0 - x)^2. \quad (9.52)$$

Для определения $\mathcal{L}_0$ используем граничное условие в точке $x = 0$. В этом случае

$$\varphi(0) = -\varphi_K = -\frac{1}{e}(\Phi_M - \Phi_n). \quad (9.53)$$

Это условие позволяет на основании (9.44) и (9.52) определить глубину проникновения контактного поля в полупроводник

$$\mathcal{L}_0 = \sqrt{2\epsilon_r\epsilon_0\varphi_K/en_0} = \sqrt{2\epsilon_r\epsilon_0(\Phi_M - \Phi_n)/e^2n_0}. \quad (9.54)$$

Из этой формулы следует, что чем меньше степень легирования полупроводника и чем больше разность работ выхода электронов из металла и полупроводника, тем больше глубина проникновения в полупроводник электрического поля, вызванного контактной разностью потенциалов. Из (8.35) и (9.54) следует, что $\mathcal{L}_0/l_s = \sqrt{2(\Phi_M - \Phi_n)/kT}$. Поэтому в случае контакта металл—полупроводник при разности работ выхода около 1 эВ глубина проникновения поля в полупроводник примерно в 10 раз превышает длину экранирования для контакта металл—вакуум—полупроводник.

Согласно рис. 9.7, а и 9.8, б приконтактный слой полупроводника, когда он является запорным, обеднен основными носителями заряда, а поэтому обладает повышенным сопротивлением по сравнению с толщиной полупроводника. По существу контакт металл—полупроводник обладает свойствами конденсатора, поскольку он включает слой повышенного сопротивления, заключенный между металлом и сравнительно хорошо проводящим полупроводником. Поэтому емкость контакта на единицу площади, называемая барьерной емкостью, равна:

$$C = \epsilon_r\epsilon_0/\mathcal{L} = \sqrt{\epsilon_r\epsilon_0en_0/2\varphi_K}. \quad (9.55)$$

9.5. ВЫПРЯМЛЕНИЕ ТОКА В КОНТАКТЕ МЕТАЛЛ-ПОЛУПРОВОДНИК

Вернемся к рассмотрению контакта металла с донорным полупроводником, когда выполняется условие $\Phi_m > \Phi_n$. В равновесном состоянии в приконтактной области имеет место изгиб зон и возникает запирающий слой, обедненный электронами (рис. 9.7, а). В глубине полупроводника ($x > \mathcal{L}_0$) концентрация свободных носителей

$$n_0 = N_c e^{-(E_c - F)/kT} = N_c e^{\xi_0/kT}, \quad (9.56)$$

где E_c — положение дна зоны проводимости в толще полупроводника; $\xi_0 = F - E_c$ — положение уровня Ферми в глубине полупроводника, отсчитываемого от нижнего края зоны проводимости. При термодинамическом равновесии положение уровня Ферми для всей структуры одинаково:

$$F = E_c + \xi_0 = E_c + e\varphi(x) + \xi(x) = \text{const}. \quad (9.57)$$

Здесь $e\varphi(x)$ — потенциальная энергия электронов в приконтактной области; $\xi(x)$ — положение уровня Ферми в приконтактной области, отсчитываемого от дна зоны проводимости.

В случае приложения к контакту металл—полупроводник внешнего напряжения состояние полупроводника становится неравновесным и концентрация электронов проводимости будет определяться квазиуровнем Ферми (см. § 7.1), равным:

$$F_n = E_c + e\varphi_1(x) + \xi_1(x), \quad (9.58)$$

где $\xi_1(x)$ и $\varphi_1(x)$ — квазиуровень Ферми, отсчитанный от дна зоны проводимости, и потенциальная энергия электронов в приконтактной области полупроводника. Поэтому концентрацию свободных электронов в приконтактной области полупроводника, т. е. в области $0 < x < \mathcal{L}$, по аналогии с (9.56) можно записать в виде

$$n(x) = N_c e^{\xi_1(x)/kT}. \quad (9.59)$$

Поскольку в приконтактной области концентрация электронов зависит от x , то квазиуровень Ферми непостоянен во всей области полупроводника. Чтобы определить его изменение, вычислим ток, протекающий через полупроводник. Плотность тока

$$J = en\mu_n \mathcal{E} + eD_n \frac{dn}{dx} \quad (9.60)$$

можно выразить через потенциал и концентрацию электронов, используя для этого (9.59). В результате получим:

$$J = n\mu_n \frac{d(e\varphi)}{dx} + \frac{eD_n}{kT} \frac{d\xi_1}{dx}. \quad (9.61)$$

Это выражение с учетом соотношения Эйнштейна преобразуется к виду

$$J = n\mu_n \frac{d(e\varphi + \xi_1)}{dx}. \quad (9.62)$$

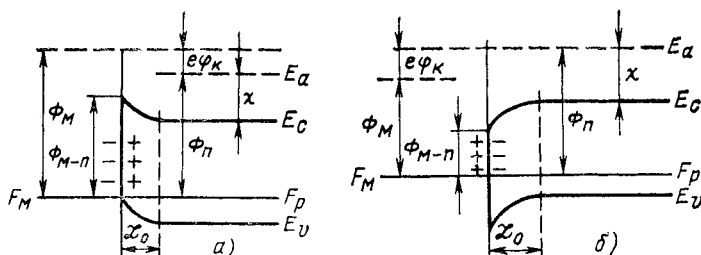


Рис. 9.10. Прямое (а) и обратное (б) смещение на контакте металл—полупроводник

Сравнивая (9.62) и (9.58), можно написать:

$$J = n\mu_n \frac{dF_n}{dx}. \quad (9.63)$$

Найдем, как изменится квазиуровень Ферми между двумя точками 1 и 2 полупроводника, в которых плотность тока одинакова:

$$\int_{(1)}^{(2)} dF_n = \int_{(1)}^{(2)} \frac{J dx}{n\mu_n} = eJs \int_{(1)}^{(2)} \frac{dx}{en\mu_n s} = eI \int_{(1)}^{(2)} dR, \quad (9.64)$$

где $dR = \frac{dx}{\sigma s}$ — сопротивление участка полупроводника длиной dx с поперечным сечением s и удельной проводимостью $\sigma = en\mu_n$.

Интегрируя уравнение (9.64), будем иметь:

$$\Delta F_n = F_{n_2} - F_{n_1} = eI (R_2 - R_1) = e(U_2 - U_1). \quad (9.65)$$

Таким образом, при наличии внешнего электрического поля изменение положения квазиуровня Ферми определяется приложенной разностью потенциалов.

Определим теперь плотность тока, текущего через контакт металл—полупроводник, при различной полярности внешнего напряжения. Предположим, что к полупроводнику приложено отрицательное напряжение, а к металлу — положительное (прямое смещение). Будем полагать, что все внешнее напряжение U приходится на приповерхностный слой полупроводника, обладающего высоким сопротивлением. В этом случае положение уровня Ферми в объеме полупроводника относительно уровня E_C останется постоянным, но по отношению к положению уровня Ферми в металле будет смещено на величину eU . Следовательно, контактная разность потенциалов между металлом и полупроводником уменьшится на U и станет равной $\phi_K - U$, а в приконтактной области положение квазиуровня Ферми изменится от F_n до F_M (рис. 9.10, а). В результате снижения потенциального барьера со стороны полупроводника равновесие в системе нарушается и потечет ток, равный разности токов термоэлектронной эмиссии, направленных от металла к полупроводнику и в обратном направлении:

$$J = J_1 - J_2. \quad (9.66)$$

Снижение потенциального барьера со стороны полупроводника вызовет более интенсивный переход электронов из полупроводника в металл по сравнению с равновесным состоянием. Этот переход электронов обусловит ток термоэлектронной эмиссии, направленный от металла к полупроводнику и равный:

$$J_1 = AT^2 e^{-(\Phi_{\text{п}} + e\Phi_{\text{к}} - eU)/kT} = AT^2 e^{-(\Phi_{\text{м}} - eU)/kT}. \quad (9.67)$$

При положительном напряжении на металле (рис. 9.10, а) высота потенциального барьера со стороны металла не изменилась, поэтому поток электронов из металла в полупроводник при неравновесном состоянии будет тот же, что и в равновесном состоянии. Он создает ток термоэлектронной эмиссии, направленный от полупроводника к металлу и равный:

$$J_2 = AT^2 e^{-\Phi_{\text{м}}/kT}. \quad (9.68)$$

Результирующая плотность тока, текущего через контакт, будет определяться выражением вида

$$J = J_1 - J_2 = AT^2 e^{-\Phi_{\text{м}}/kT} (e^{eU/kT} - 1) = J_s (e^{eU/kT} - 1). \quad (9.69)$$

Величина

$$J_s = AT^2 e^{-\Phi_{\text{м}}/kT} \quad (9.70)$$

носит название плотности тока насыщения.

В случае отрицательного напряжения на металле (обратное смещение) в приконтактной области полупроводника потенциальный барьер увеличится на приложенную разность потенциалов и выражение для плотности тока, текущего через контакт, примет вид:

$$J = AT^2 e^{-(\Phi_{\text{п}} + e\Phi_{\text{к}} + eU)/kT} - AT^2 e^{-\Phi_{\text{м}}/kT} = J_s (e^{-eU/kT} - 1). \quad (9.71)$$

Если при прямом смещении считать $U > 0$, а при обратном смещении $U < 0$, равенства (9.69) и (9.71) можно записать как

$$J = J_s (e^{eU/kT} - 1). \quad (9.72)$$

Из (9.72) следует, что при прямом смещении ($U > 0$) плотность тока через контакт в зависимости от напряжения растет по экспоненциальному закону, а при обратном смещении ($U < 0$) плотность тока, возрастая, стремится к J_s . Таким образом, контакт металл—полупроводник обладает выпрямляющим действием и ведет себя как выпрямитель переменного тока.

Область полупроводника вблизи контакта представляет собой слой объемного заряда, толщина которого при отсутствии внешнего напряжения выражается формулой (9.54). В случае подачи на контакт внешнего напряжения толщина слоя объемного заряда

$$\mathcal{L} = \sqrt{\frac{2\epsilon_r \epsilon_0 (\Phi_{\text{к}} - U)}{e n_0}}, \quad (9.73)$$

где, как и в равенстве (9.72), знак U зависит от полярности приложенного напряжения.

При прямом смещении ($U > 0$) толщина слоя $\mathcal{L}_{np}$ объемного заряда меньше по сравнению с толщиной запорного слоя $\mathcal{L}_0$ при термодинамическом равновесии, а в случае обратного смещения ($U < 0$) $\mathcal{L}_{обp} > \mathcal{L}_0$.

9.6. ДИОДНАЯ ТЕОРИЯ ВЫПРЯМЛЕНИЯ ТОКА

Расчет тока насыщения вольт-амперной характеристики контакта металл—полупроводник можно провести на основе диодной и диффузионной теорий.

Если длина свободного пробега электронов l значительно меньше толщины запорного слоя $\mathcal{L}_0$, носители заряда, проходя через этот слой, многократно рассеиваются. В этом случае необходимо учитывать как диффузионную, так и дрейфовую составляющие тока. Использование закона Ома для нахождения дрейфового тока возможно в том случае, если электрическое поле слабое, условием чего согласно (6.184) является:

$$\mathcal{E}_s \ll \delta \frac{kT}{el}. \quad (9.74)$$

Максимальная напряженность поля $\mathcal{E}_s$ будет на границе полупроводника с металлом ($x = 0$). С учетом (9.52) и (9.54)

$$\mathcal{E}_s = \mathcal{E}_{max} = - \left. \frac{d\varphi}{dx} \right|_{x=0} = - \frac{en_0 \mathcal{L}_0}{\epsilon_r \epsilon_0} = - \frac{2\varphi_R}{\mathcal{L}_0}. \quad (9.75)$$

Подставляя найденное значение $\mathcal{E}_s$ в (9.74), получаем условие применимости диффузионной теории выпрямления:

$$\mathcal{L}_0 \gg \frac{2e\varphi_R}{\delta kT} l. \quad (9.76)$$

Если соотношение длины свободного пробега электронов l и толщины запорного слоя $\mathcal{L}_0$ таково, что выполняется обратное неравенство

$$l \gg \frac{\delta kT}{2e\varphi_R} \mathcal{L}_0, \quad (9.77)$$

то можно считать, что носители заряда пролетают запорную область, почти не испытывая столкновений. В этом случае запирающий слой объемного заряда полупроводника подобен вакуумному промежутку между электродами электронной лампы. В связи с этим теория, используемая в этом случае для расчета вольт-амперной характеристики контакта металл—полупроводник, получила название диодной.

Согласно (9.77) диодная теория применима в том случае, если

$$\frac{\delta kT \mathcal{L}_0}{2e\varphi_R l} \ll 1. \quad (9.78)$$

Это условие будет выполняться тем лучше, чем меньше $\mathcal{L}_0$ и больше l . Но толщина запорного слоя, как следует из (9.54), тем меньше, чем больше равновесная концентрация свободных электронов n_0 , а длина свободного пробега электронов l согласно (6.29) тем больше, чем больше их подвижность. Следовательно, диодная

теория выпрямления тока в контакте металл-полупроводник применима к полупроводникам, обладающим большой концентрацией и высокой подвижностью носителей заряда.

Например, для германия электронного типа $\mu_n = 3900 \text{ см}^2/(\text{В} \cdot \text{с})$; $\epsilon_r = 16,5$; $m_n^* = 0,3m_0$. При $n_0 = 10^{14} \text{ см}^{-3}$ $e\varphi_K = 0,3 \text{ эВ}$, $\mathcal{L}_0 = 2,3 \cdot 10^{-4} \text{ см}$, а $l = 1,5 \cdot 10^{-5} \text{ см}$ и отношение $kT\mathcal{L}_0/2e\varphi_K l = 0,6 < 1$. Если же повысить концентрацию электронов до $n_0 = 10^{16} \text{ см}^{-3}$, то $kT\mathcal{L}_0/2e\varphi_K l = 0,06$. Применение диодной теории выпрямления в этом случае полностью оправдано.

В полупроводниках с малым значением концентрации и подвижности носителей заряда дело обстоит по-иному. Например, в закиси меди p -типа $\mu_p = 60 \text{ см}^2/(\text{В} \cdot \text{с})$; $\epsilon_r = 8,75$. При $n_0 = 10^{14} \text{ см}^{-3}$ $e\varphi_K = 0,5 \text{ эВ}$; $\mathcal{L}_0 = 2,2 \cdot 10^{-4} \text{ см}$, а $l = 4 \cdot 10^{-7} \text{ см}$ и $kT\mathcal{L}_0/2e\varphi_K l = 15 \gg 1$. В данном случае для расчета вольт-амперной характеристики (ВАХ) контакта металл-полупроводник должна применяться диффузионная теория выпрямления.

При рассмотрении контакта металла с полупроводником (§ 9.4) считалось, что из металла способны выйти лишь электроны с энергией, превышающей значение Φ_M . Однако при обычных температурах число электронов, обладающих столь высокой энергией, исчезающе мало. Но при наличии очень тонкого потенциального барьера между образцами электроны могут преодолевать его в результате туннельного эффекта.

Как следует из квантовой механики, прозрачность D потенциального барьера высотой Φ_0 и толщиной d выражается формулой

$$D = e^{-\frac{2d}{\hbar} [2m^* (\Phi_0 - E_x)]^{1/2}}. \quad (9.79)$$

Здесь $E_x = \frac{m^* v_x^2}{2}$; v_x — скорость электронов в направлении к поверхности образца. Если принять, что $\Phi_0 \approx 1 \text{ эВ}$; $d = 10^{-8} \text{ см}$, то при условии $E_x \ll \Phi_0$ получим $D = 1/3$. Это означает, что столь тонкие барьеры практически прозрачны для электронов. Поэтому в дальнейшем будем полагать, что все электроны, энергия которых превышает величину Φ_{M-n} , способны перейти в полупроводник.

Для перехода из полупроводника в металл электроны должны иметь в направлении x достаточную скорость, чтобы преодолеть потенциальный барьер высотой $e(\varphi_K - U)$. Ток, обусловленный такими электронами, согласно (9.23) будет выражаться соотношением

$$J_1 = e \int_{v_{x \min}}^{\infty} \int_{-\infty}^{\infty} v_x dn = \frac{4\pi e m^* (kT)^2}{h^3} e^{(F - E_c)/kT} e^{-(\varphi_K - U)/kT}. \quad (9.80)$$

Подставляя сюда выражения для равновесной концентрации электронов в толще полупроводника

$$n_0 = 2 \left(2\pi m kT/h^2 \right)^{3/2} e^{(F - E_c)/kT} \quad (9.81)$$

и их средней тепловой скорости

$$v_0 = (8kT/\pi m^*)^{1/2}, \quad (9.82)$$

получаем:

$$J_1 = \frac{1}{4} en_0 v_0 e^{-e(\varphi_K - U)/kT} = \frac{1}{4} en_s v_0 e^{eU/kT}. \quad (9.83)$$

Здесь

$$n_s = n_0 e^{-e\varphi_K/kT} \quad (9.84)$$

— концентрация электронов на поверхности полупроводника при $U = 0$.

Со стороны металла высота потенциального барьера остается та же самая (см. рис. 9.10), поэтому поток электронов из металла в полупроводник не зависит от приложенного напряжения и равен потоку электронов из полупроводника в состоянии термодинамического равновесия

$$J_2 = J_1|_{U=0} = en_s v_0 / 4. \quad (9.85)$$

Плотность суммарного тока через контакт будет определяться выражением

$$J = J_1 - J_2 = \frac{1}{4} en_s v_0 (e^{eU/kT} - 1) = J_s (e^{eU/kT} - 1), \quad (9.86)$$

где

$$J_s = \frac{1}{4} en_s v_0 \quad (9.87)$$

есть ток насыщения для диодной теории выпрямления, который, как это следует из (9.87), определяется тепловой скоростью электронов и их концентрацией у поверхности полупроводника.

9.7. ДИФFUЗИОННАЯ ТЕОРИЯ ВЫПРЯМЛЕНИЯ ТОКА

В диффузионной теории выпрямления при определении тока, текущего через контакт металл—полупроводник, учитываются как дрейфовая, так и диффузионная составляющие тока. В этом случае для расчета ВАХ поступим следующим образом. Умножим обе части (9.60) на $e^{e\varphi(x)/kT}$. Используя соотношение Эйнштейна и выражая напряженность поля через потенциал, будем иметь:

$$\begin{aligned} J e^{e\varphi(x)/kT} &= n \mu_n e^{e\varphi(x)/kT} \frac{d(e\varphi)}{dx} + \mu_n kT e^{e\varphi(x)/kT} \frac{dn}{dx} = \\ &= \mu_n kT \frac{d}{dx} [n e^{e\varphi(x)/kT}]. \end{aligned} \quad (9.88)$$

Проинтегрируем обе части (9.88) по x в пределах от 0 до $\mathcal{L}$:

$$J \int_0^{\mathcal{L}} e^{e\varphi(x)/kT} dx = \mu_n kT \int_0^{\mathcal{L}} \frac{d}{dx} [n e^{e\varphi(x)/kT}] dx. \quad (9.89)$$

Выразим отсюда плотность тока J , которая не зависит от x :

$$J = \frac{\mu_n kT [n(\mathcal{L}) e^{e\varphi(\mathcal{L})/kT} - n(0) e^{e\varphi(0)/kT}]}{\int_0^{\mathcal{L}} e^{e\varphi(x)/kT} dx}. \quad (9.90)$$

Для вычисления числителя (9.90) используем граничные условия. В объеме полупроводника поле отсутствует, поэтому

$$\varphi(\mathcal{L}) = 0; n(\mathcal{L}) = n_0, \quad (9.91)$$

а на границе раздела с металлом ($x = 0$)

$$\varphi(0) = \varphi_K - U. \quad (9.92)$$

Рассмотрим случай, когда внешнее напряжение сравнительно невелико и в приконтактной области полупроводника всегда имеется запиорный слой

$$(\varphi_K - U) \gg kT/e. \quad (9.93)$$

Можно показать, что при этом условии концентрация электронов на поверхности полупроводника практически не зависит от напряжения, так что можно без большой ошибки положить:

$$n(0) = n_s = n_0 e^{-\varepsilon \varphi_K / kT}. \quad (9.94)$$

С учетом граничных условий (9.91), (9.92) и (9.94) выражение (9.90) для плотности тока принимает вид

$$J = \frac{n_0 \mu_n kT (1 - e^{-\varepsilon U / kT})}{\int_0^{\mathcal{L}} e^{\varepsilon \varphi(x) / kT} dx}. \quad (9.95)$$

Для вычисления знаменателя дроби (9.95) произведем замену переменной интегрирования:

$$\int_0^{\mathcal{L}} e^{\varepsilon \varphi(x) / kT} dx = \int_{\varphi_K - U}^0 e^{\varepsilon \varphi(x) / kT} \left(\frac{d\varphi}{dx} \right)^{-1} d\varphi. \quad (9.96)$$

Наибольший вклад в интеграл будет вносить та область, где значение $\varphi(x)$ максимально, т. е. участок вблизи $x = 0$. С увеличением x функция, стоящая под знаком интеграла, в результате уменьшения $\varphi(x)$ будет резко падать и вкладом ее можно пренебречь. Поэтому для приближенного вычисления интеграла можно вынести за знак интеграла значение $(d\varphi/dx)^{-1}$ в точке $x = 0$:

$$\left(\frac{d\varphi}{dx} \right)^{-1} = -\mathcal{E}_s^{-1}. \quad (9.97)$$

На основании (9.97) будем иметь:

$$\begin{aligned} \int_{\varphi_K - U}^0 e^{\varepsilon \varphi(x) / kT} \left(\frac{d\varphi}{dx} \right)^{-1} d\varphi &= -\mathcal{E}_s^{-1} \frac{kT}{e} [1 - e^{\varepsilon (\varphi_K - U) / kT}] = \\ &= \mathcal{E}_s^{-1} \frac{kT}{e} e^{\varepsilon (\varphi_K - U) / kT}. \end{aligned} \quad (9.98)$$

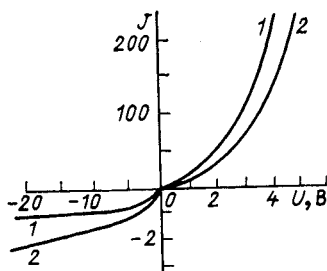


Рис. 9.11. Вольт-амперные характеристики контакта металл—полупроводник:

1 — диодная теория; 2 — диффузионная теория (в относительных единицах)

Здесь мы пренебрегли единицей, стоящей в скобках, так как в соответствии с условием (9.93)

$$e^{e(\varphi_K - U)/kT} \gg 1. \quad (9.99)$$

Подставляя в (9.95) выражение (9.98), находим:

$$J = en_0 \mu_n \mathcal{E}_s e^{-e\varphi_K/kT} (e^{eU/kT} - 1). \quad (9.100)$$

Используя (9.94), получаем выражение вольт-амперной характеристики контакта металл—полупроводник для диффузионной теории выпрямления:

$$\begin{aligned} J &= en_s \mu_n \mathcal{E}_s (e^{eU/kT} - 1) = \\ &= J_s (e^{eU/kT} - 1). \end{aligned} \quad (9.101)$$

Здесь величина $J_s = en_s \mu_n \mathcal{E}_s = en_s v_s$ др, определяющая значение тока при больших обратных смещениях, в отличие от плотности тока насыщения в диодной теории зависит от приложенного напряжения, т. е. выражается через дрейфовую скорость электрона на поверхности полупроводника.

ВАХ контакта металл—полупроводник для диодной и диффузионной теорий приведены на рис. 9.11 (кривые 1 и 2 соответственно).

9.8. КОНТАКТ ЭЛЕКТРОННОГО И ДЫРОЧНОГО ПОЛУПРОВОДНИКОВ

Рассмотрим явления, возникающие на контакте электронного и дырочного полупроводников. Для создания такого контакта в полупроводник вводятся как донорная, так и акцепторная примеси. При этом концентрации доноров и акцепторов меняются так, что в одной части образец содержит доноры и обладает электронной электропроводностью, а в другой части содержит акцепторы и обладает дырочной электропроводностью, а следовательно, в некоторой области кристалла происходит смена электропроводности с электронной на дырочную. Такой переход между материалами с электропроводностью n - и p -типа носит название p - n перехода.

Предположим, что переход бесконечно узкий и акцепторная область полупроводника легирована сильнее, чем электронная, т. е. $N_a > N_d$ (рис. 9.12, а). Пусть в p -области концентрация основных носителей заряда — дырок p_p и неосновных носителей заряда — электронов n_p , а в n -области концентрация электронов n_n и дырок p_n . Предположим также, что энергетические уровни доноров и акцепторов расположены так близко к зоне проводимости и валентной зоне, что все они полностью ионизованы. Тогда $p_p = N_a$, $n_n = N_d$ и для состояния термодинамического равновесия в случае отсутствия вырождения

$$p_p n_p = n_n p_n = n_i^2. \quad (9.102)$$

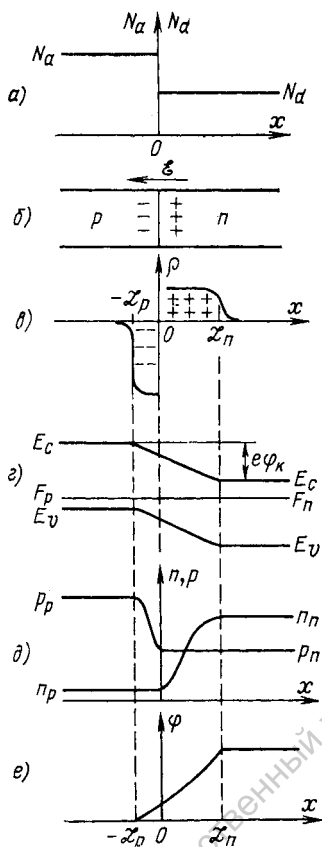


Рис. 9.12. Распределение примесей (а), разделение зарядов и возникновение электрического поля (б), распределение объемного заряда (в), зонная структура (г), распределение концентрации электронов и дырок (д) и изменение потенциала (е) в контакте электронного и дырочного полупроводников

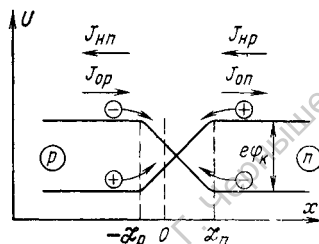


Рис. 9.13. Потенциальные барьеры для дырок и электронов в p - n переходе

В первый момент соприкосновения n - и p -областей вблизи границы перехода будет существовать большой градиент концентрации электронов и дырок. В результате начнется диффузия электронов из n -области в p -область и дырок из p -области в n -область. Возникновение диффузионных потоков приведет к разделению зарядов,

вследствие чего появится положительный объемный заряд в n -области, примыкающей к переходу, в которой заряд обусловлен положительными ионами донорной примеси, и отрицательный — в p -области около перехода, созданной отрицательными ионами акцепторной примеси (рис. 9.12, б, в). Эти объемные заряды в области контакта создадут сильное электрическое поле, направленное от n -области к p -области и препятствующее движению электронов и дырок. В результате установится равновесное состояние, которое будет характеризоваться постоянством уровня Ферми для всего полупроводника, а в области перехода, где имеется электрическое поле, зоны энергии будут искривлены (рис. 9.12, г). Искривление зон энергии вызовет перераспределение концентрации электронов и дырок (рис. 9.12, д) и изменит ход электростатического потенциала в области p - n перехода (рис. 9.12, е).

Как следует из рис. 9.13, основные носители заряда при переходе через контакт должны преодолевать потенциальный барьер высотой $e\varphi_k$. Переход неосновных носителей заряда совершается под дей-

ствием электрического поля p - n перехода. В состоянии термодинамического равновесия плотность диффузионного тока основных носителей заряда J_{0p} и J_{0n} уравновешена плотностью дрейфового тока неосновных носителей заряда J_{np} и J_{nn} и суммарный ток через p - n переход равен нулю.

Предположим, что Φ_n есть термоэлектронная работа выхода электронного полупроводника, а Φ_p — дырочного. Тогда потенциальный барьер на p - n переходе при термодинамическом равновесии определяется уравнением

$$e\varphi_K = \Phi_p - \Phi_n = (\chi - E_c + F_p) - (\chi - E_c + F_n) = F_p - F_n. \quad (9.103)$$

Так как донорные и акцепторные примеси полностью ионизованы, то, используя (9.27) и (9.29) или (4.127) и (4.136), получаем:

$$e\varphi_K = E_g - kT \ln \frac{N_D}{N_A} - kT \ln \frac{N_c}{N_d}. \quad (9.104)$$

Учитывая, что

$$n_i^2 = N_c N_v e^{-E_g/kT}, \quad (9.105)$$

из (9.104) получаем:

$$e\varphi_K = kT \ln \frac{n_n p_p}{n_i^2} \quad (9.106)$$

или

$$\frac{p_n}{p_p} = \frac{n_p}{n_n} = e^{-e\varphi_K/kT}. \quad (9.107)$$

Таким образом, контактная разность потенциалов на p - n переходе тем больше, чем сильнее легированы n и p -области полупроводника; ее максимальное значение для невырожденного полупроводника равно:

$$\varphi_{K \max} = \frac{1}{e} E_g. \quad (9.108)$$

На основании (9.107) можно получить следующие соотношения для концентрации основных и неосновных носителей заряда:

$$p_n = p_p e^{-e\varphi_K/kT}; \quad n_p = n_n e^{-e\varphi_K/kT}. \quad (9.109)$$

Работа выхода электронов из полупроводника довольно велика (несколько электрон-вольт), и при комнатной температуре практически нет электронов, обладающих достаточной энергией, чтобы покинуть кристалл. Однако электроны вполне могут переходить из n -области непосредственно в p -область, преодолевая возникший потенциальный барьер. Например, для германия при $T = 300$ К и значениях концентрации $N_d = n_n = 10^{16} \text{ см}^{-3}$; $p_n = 10^{10} \text{ см}^{-3}$; $N_a = p_p = 10^{14} \text{ см}^{-3}$; $n_p = 10^{12} \text{ см}^{-3}$ потенциальный барьер составляет:

$$e\varphi_K = kT \ln \frac{p_p}{p_n} = 0,25 \text{ эВ.}$$

Поскольку p -область легирована сильнее, чем n -область, то $p_p \gg n_n$ и, следовательно, толщина слоя объемного заряда низкоом-

ной дырочной области $\mathcal{L}_p$ будет меньше, чем электронной $\mathcal{L}_n$, т. е. $\mathcal{L}_n > |\mathcal{L}_p|$ (рис. 9.12, в), а вся область объемного заряда составит:

$$\mathcal{L}_0 = \mathcal{L}_n + \mathcal{L}_p. \quad (9.110)$$

В интервале $-\mathcal{L}_p \leq x < 0$ объемный заряд отрицательный, определяется концентрацией ионов акцепторной примеси:

$$\rho = -eN_a^- = -ep_p \quad (9.111)$$

и уравнение Пуассона для этой области будет иметь вид:

$$\frac{d^2\varphi}{dx^2} = \frac{ep_p}{\varepsilon_r \varepsilon_0}. \quad (9.112)$$

В интервале $0 < x < \mathcal{L}_n$ объемный заряд положительный, определяется ионами донорной примеси:

$$\rho = eN_d^+ = en_n \quad (9.113)$$

и уравнение Пуассона запишется в форме

$$\frac{d^2\varphi}{dx^2} = -\frac{en_n}{\varepsilon_r \varepsilon_0}. \quad (9.114)$$

Граничными условиями в этом случае являются:

$$\varphi(-\mathcal{L}_p) = 0; \quad \left. \frac{d\varphi}{dx} \right|_{x=-\mathcal{L}_p} = 0; \quad (9.115)$$

$$\varphi(\mathcal{L}_n) = \varphi_K; \quad \left. \frac{d\varphi}{dx} \right|_{x=\mathcal{L}_n} = 0. \quad (9.116)$$

Этим условиям удовлетворяют следующие выражения для φ , являющиеся согласно (9.52) решением уравнений (9.112) и (9.114):
при $-\mathcal{L}_p \leq x < 0$

$$\varphi_p = \frac{ep_p}{2\varepsilon_r \varepsilon_0} (\mathcal{L}_p + x)^2; \quad (9.117)$$

при $0 < x \leq \mathcal{L}_n$

$$\varphi_n = \varphi_K - \frac{en_n}{2\varepsilon_r \varepsilon_0} (\mathcal{L}_n - x)^2. \quad (9.118)$$

При $x = 0$ потенциал (рис. 9.12, е) и его производная непрерывны, поэтому

$$\varphi_p(0) = \varphi_n(0); \quad (9.119)$$

$$\left. \frac{d\varphi_p}{dx} \right|_{x=0} = \left. \frac{d\varphi_n}{dx} \right|_{x=0} \quad (9.120)$$

Применяя условие (9.120) к выражениям (9.117) и (9.118), получаем:

$$n_n \mathcal{L}_n = p_p \mathcal{L}_p. \quad (9.121)$$

Следовательно, в обеих областях полупроводника, прилегающих к р-п переходу, объемные заряды равны. Это является условием сохранения электронейтральности.

Из (9.121) нетрудно получить следующие соотношения:

$$\frac{\mathcal{L}_n}{\mathcal{L}_0} = \frac{\gamma_p}{n_n + p_p}; \quad \frac{\mathcal{L}_p}{\mathcal{L}_0} = \frac{n_n}{n_n + p_p}. \quad (9.122)$$

Используя эти условия, на основании равенств (9.117)—(9.119) получаем:

$$\varphi_k = \frac{e}{2\varepsilon_r \varepsilon_0} (n_n \mathcal{L}_n^2 + p_p \mathcal{L}_p^2) = \frac{e}{2\varepsilon_r \varepsilon_0} \mathcal{L}_0^2 \frac{n_n p_p}{n_n + p_p}. \quad (9.123)$$

Отсюда следует выражение для полной толщины слоя объемного заряда p - n -перехода:

$$\mathcal{L}_0 = \sqrt{\frac{2\varepsilon_r \varepsilon_0}{e} \varphi_k \frac{n_n + p_p}{n_n p_p}}. \quad (9.124)$$

Из этой формулы видно, что чем выше степень легирования n - и p -областей полупроводника, тем меньше толщина области объемного заряда $\mathcal{L}_0$. Если одна из областей легирована значительно сильнее другой, то большая часть падения электростатического потенциала приходится на высокоомную область (рис. 9.12, е).

Как следует из рис. 9.12, д, в области p - n перехода имеет место значительное уменьшение концентрации носителей заряда. В силу этого сопротивление перехода велико по сравнению с сопротивлением слоя полупроводника n - или p -типа той же площади и толщины, что и область объемного заряда. Следовательно, *электронно-дырочный переход представляет собой слой низкой удельной проводимости, заключенный между областями высокой удельной проводимости, и обладает свойствами конденсатора*. Емкость на единицу площади, так называемая *барьерная емкость*, может быть определена по формуле

$$C = \frac{\varepsilon_r \varepsilon_0}{\mathcal{L}_0} = \sqrt{\frac{\varepsilon_r \varepsilon_0}{2\varphi_k} \frac{n_n p_p}{n_n + p_p}}. \quad (9.125)$$

9.9. ВЫПРЯМЛЕНИЕ ТОКА В p - n ПЕРЕХОДЕ

Рассмотрим теперь p - n переход, к которому приложена разность потенциалов U такая, что p -область заряжается положительно (прямое смещение) (рис. 9.14, б). Поскольку сопротивление слоя объемного заряда перехода высокое, то падение напряжения будет в основном в этой области. Поэтому при прямом смещении высота потенциального барьера понижается на eU по сравнению с равновесным состоянием (рис. 9.14, а), соответственно изменится и толщина запирающего слоя. Для ее вычисления нужно вместо формулы (9.124) использовать выражение

$$\mathcal{L}_{np} = \sqrt{\frac{2\varepsilon_r \varepsilon_0 (\varphi_k - U)}{e} \frac{n_n + p_p}{n_n p_p}}. \quad (9.126)$$

Понижение потенциального барьера приведет к увеличению потока основных носителей заряда по сравнению с равновесным состоянием, так как большее количество электронов из n -области будет переходить в p -область и большее количество дырок из p -области в n -об-

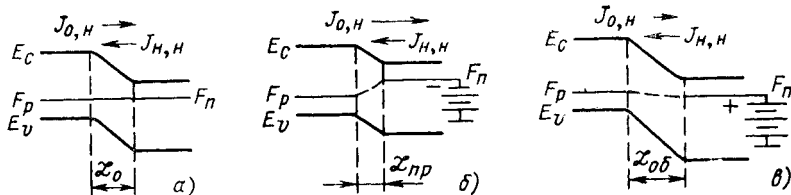


Рис. 9.14. Энергетическая диаграмма p - n перехода при термодинамическом равновесии (а), при подаче на переход прямого (б) и обратного (в) смещений

ласть, а поток неосновных носителей заряда через переход останется практически неизменным. В результате этого во внешней цепи будет протекать ток («прямой ток»), равный разности токов основных и неосновных носителей заряда и направленный от p -области к n -области (рис. 9.14, б), т. е.

$$J = J_{0,n} - J_{n,n}. \quad (9.127)$$

В n -области появившиеся избыточные неосновные носители заряда — дырки Δp создадут в первый момент вблизи контакта положительный объемный заряд. Однако через очень короткое время, определяемое максвелловским временем релаксации (см. § 7.3), этот заряд будет скомпенсирован объемным зарядом основных носителей заряда — электронов, которые под действием электрического поля, созданного избыточными дырками, будут подняты в количестве Δn из глубины n -области, а в n -область электроны поступят из внешней цепи. Таким образом, во всех частях электронного полупроводника будет соблюдаться электронейтральность, но в приконтактной области к p - n переходу концентрация электронов и дырок будет повышена на $\Delta n = \Delta p$ по сравнению с равновесным состоянием. Введение в полупроводник носителей заряда с помощью p - n перехода при подаче на него прямого смещения в область, где эти носители заряда являются неосновными, называют инжекцией. Теперь концентрация дырок в n -области вблизи контакта будет равна:

$$p = p_n + \Delta p. \quad (9.128)$$

Для ее нахождения в стационарном случае при $x = \mathcal{L}_n$ и отсутствии вырождения носителей заряда нужно в (9.109) вместо $e\varphi_K$ использовать значение $e(\varphi_K - U)$. Следовательно,

$$p(\mathcal{L}_n) = p_p e^{-e(\varphi_K - U)/kT} = p_n e^{eU/kT}, \quad (9.129)$$

откуда следует, что концентрация избыточных дырок в n -области при $x = \mathcal{L}_n$ равна:

$$\Delta p(\mathcal{L}_n) = p - p_n = p_n (e^{eU/kT} - 1). \quad (9.130)$$

Аналогичные явления происходят и в p -области: сюда из n -области инжектируются электроны и концентрация избыточных электронов при $x = -\mathcal{L}_p$ составит:

$$\Delta n(-\mathcal{L}_p) = n_p (e^{eU/kT} - 1), \quad (9.131)$$

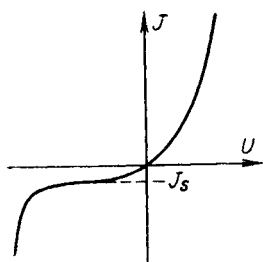


Рис. 9.15. Вольт-амперная характеристика p - n перехода

Из (9.130) и (9.131) следует, что с увеличением прямого смещения на p - n переходе концентрация инжектируемых неосновных носителей заряда резко возрастает, что приводит к сильному росту тока через контакт в прямом направлении.

Если к p - n переходу приложено обратное смещение (рис. 9.14, в), p -область заряжена отрицательно, потенциальный барьер повышается на величину eU и увеличивается толщина заборного слоя объемного заряда. $\mathcal{L}_{обр}$ вычисляется по (9.124) при подстановке значения $\varphi_n + U$. Чем сильнее смещен переход в обратном направлении, тем меньшее количество основных носителей заряда способно преодолеть возросший потенциальный барьер. В соответствии с этим количество неосновных носителей заряда в приконтактной области уменьшается по сравнению с равновесным состоянием, следовательно, уменьшается и количество основных носителей заряда вследствие соблюдения электронейтральности. Это явление носит название *экстракции носителей заряда*. Избыточная концентрация электронов в p -области по-прежнему также будет определяться равенством (9.131), но значение U теперь берется с отрицательным знаком. Таким образом, при обратном смещении p - n перехода ток основных носителей заряда будет меньше, чем при равновесном состоянии, а ток неосновных носителей заряда практически не изменяется. Поэтому суммарный ток через p - n переход будет направлен от n -области к p -области и с увеличением обратного напряжения вначале будет незначительно расти, а затем стремиться к некоторой величине, называемой током насыщения. Следовательно, p - n переход имеет нелинейную ВАХ (рис. 9.15).

9.10. ТЕОРИЯ ТОНКОГО p - n ПЕРЕХОДА

Проведем расчет ВАХ p - n перехода. Для упрощения решения задачи примем следующие допущения.

1. Переход такой тонкий, что носители заряда пролетают через слой объемного заряда без рекомбинации. Это означает, что толщина заборного слоя значительно меньше диффузионной длины:

$$\mathcal{L}_0 \ll L_D. \quad (9.132)$$

2. Обе области полупроводника сильно легированы, т. е. $p_p \gg n_i$ и $n_n \gg n_i$. Благодаря этому падением напряжения на образце можно пренебречь.

3. Контакты полупроводника с металлом, к которым приложена внешняя разность потенциалов, сделаны антизапорными и находятся далеко от перехода, так что неравновесные носители заряда рекомбинируют полностью, не доходя до них. Падением напряжения

на этих контактах также можно пренебречь и считать, что вся внешняя разность потенциалов приложена к p - n переходу.

4. Центры захвата и поверхностная рекомбинация отсутствуют, и уменьшение избыточной концентрации носителей заряда происходит только в результате рекомбинации их в объеме полупроводника, которую будем считать линейной.

Чтобы рассчитать ВАХ p - n перехода, нужно найти закон изменения концентрации свободных носителей заряда в p - и n -областях. Для этого необходимо решить уравнения непрерывности для дырок и электронов, которые имеют вид:

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial J_p}{\partial x} - \frac{\Delta p}{\tau_p}; \quad (9.133)$$

$$\frac{\partial n}{\partial t} = \frac{1}{e} \frac{\partial J_n}{\partial x} - \frac{\Delta n}{\tau_n}. \quad (9.134)$$

Будем считать, что в p - и n -областях значения коэффициентов и подвижности дырок и электронов соответственно одинаковы (что, строго говоря, не совсем верно). Плотность полного тока дырок и электронов, определяющаяся дрейфовыми и диффузионными составляющими, будет равна:

$$\left. \begin{aligned} J_p &= e p \mu_p \mathcal{E} - e D_p dp/dx; \\ J_n &= e n \mu_n \mathcal{E} + e D_n dn/dx, \end{aligned} \right\} \quad (9.135)$$

где $\mathcal{E}$ — напряженность внешнего электрического поля.

Рассмотрим n -область. При прямом смещении p - n перехода концентрация неравновесных электронов $n = n_n + \Delta n$. Здесь $\Delta n = \Delta p$ — количество электронов, подошедших к контакту из глубины n -области для нейтрализации заряда дырок, инжектированных из p -области. Поскольку n -область сильно легирована, то $n_n \gg \Delta n$ и $n \simeq n_n$, поэтому дрейфовая составляющая тока $J_{n \text{ др}}^{(n)}$ значительно превосходит диффузионную, т. е. $J_{n \text{ др}}^{(n)} \gg J_{n \text{ диф}}^{(n)}$. Следовательно, в n -области плотность тока электронов $J_n^{(n)}$ примерно будет равна дрейфовой составляющей тока электронов:

$$J_n^{(n)} = J_{n \text{ др}}^{(n)} = e n \mu_n \mathcal{E}. \quad (9.136)$$

Но в n -области $n_n \gg p_n$, $\Delta p \gg p_n$ и концентрация неравновесных дырок вблизи p - n перехода, равная $p = p_n + \Delta p$, в основном определяется количеством избыточных дырок Δp , инжектированных из p -области. Поэтому диффузионная составляющая дырочного тока превышает дрейфовую, т. е. $J_{p \text{ диф}}^{(n)} \gg J_{p \text{ др}}^{(n)}$, и

$$J_p^{(n)} = J_{p \text{ диф}}^{(n)} = -e D_p \frac{dp}{dx}. \quad (9.137)$$

С учетом (9.137) уравнение непрерывности для дырок в n -области для стационарного случая ($\frac{\partial p}{\partial t} = 0$) запишется в виде

$$D_p \frac{d^2 p}{dx^2} - \frac{p - p_n}{\tau_p} = 0. \quad (9.138)$$

Используя соотношение $L_p^2 = D_p \tau_p$, получаем:

$$\frac{d^2 \Delta p}{dx^2} - \frac{\Delta p}{L_p^2} = 0. \quad (9.139)$$

Общее решение такого уравнения имеет вид:

$$\Delta p = A e^{-x/L_p} + B e^{x/L_p}. \quad (9.140)$$

Так как при движении избыточных дырок в объем полупроводника ($x \rightarrow \infty$) концентрация их спадает до нуля, константа B должна быть принята равной нулю. Тогда

$$p = p_n + \Delta p = p_n + A e^{-x/L_p}. \quad (9.141)$$

Концентрация неравновесных дырок на границе запорного слоя при $x = \mathcal{L}_n$ согласно (9.129) равна:

$$p(\mathcal{L}_n) = p_n e^{eU/kT}. \quad (9.142)$$

Учитывая это, из (9.141) при $x = \mathcal{L}_n$ находим:

$$A = p_n (e^{eU/kT} - 1) e^{\mathcal{L}_n/L_p}. \quad (9.143)$$

Окончательно закон изменения концентрации неравновесных дырок в n -области при $x > \mathcal{L}_a$ принимает вид:

$$p(x) = p_n + p_n (e^{eU/kT} - 1) e^{-(x - \mathcal{L}_n)/L_p}, \quad (9.144)$$

и для дырочного тока на основании (9.137) получаем:

$$J_p^{(n)} = \frac{e D_p p_n}{L_p} (e^{eU/kT} - 1) e^{-(x - \mathcal{L}_n)/L_p}. \quad (9.145)$$

Проведя аналогичные расчеты, найдем, что в p -области изменение концентрации неравновесных электронов при $x < -\mathcal{L}_p$ определяется соотношением

$$n(x) = n_p + n_p (e^{eU/kT} - 1) e^{(x + \mathcal{L}_p)/L_n}. \quad (9.146)$$

а электронная составляющая тока имеет вид:

$$J_n^{(p)} = \frac{e D_n n_p}{L_n} (e^{eU/kT} - 1) e^{(x + \mathcal{L}_p)/L_n}. \quad (9.147)$$

В любом сечении полупроводника сумма плотностей электронного и дырочного токов постоянна:

$$J = J_p^{(p)} + J_n^{(p)} = J_n^{(n)} + J_p^{(n)} = \text{const}. \quad (9.148)$$

Поскольку слой объемного заряда достаточно узок и внутри его нет рекомбинации носителей заряда, то дырочные токи в p - и n -областях на границе запорного слоя одинаковы:

$$J_p^{(p)}|_{x=-\mathcal{L}_p} = J_p^{(n)}|_{x=\mathcal{L}_n}. \quad (9.149)$$

Принимая это во внимание, для плотности общего тока, текущего через p - n переход, можно использовать формулу

$$J = J_p^{(p)}|_{x=-\mathcal{L}_p} + J_n^{(p)}|_{x=-\mathcal{L}_p} = J_p^{(n)}|_{x=\mathcal{L}_n} + J_n^{(p)}|_{x=-\mathcal{L}_p}. \quad (9.150)$$

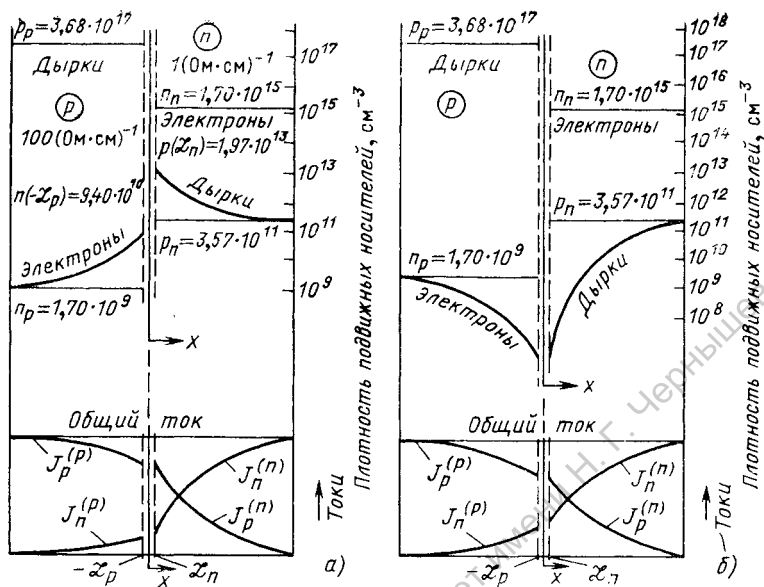


Рис. 9.16. Концентрация носителей заряда и токи в p - n переходе при прямом (а) и обратном (б) смещении

Плотности токов на основании (9.145) и (9.147) соответственно равны:

$$J_p^{(n)}|_{x=\mathcal{L}_n} = \frac{eD_p p_n}{L_p} (e^{eU/kT} - 1); \quad (9.151)$$

$$J_n^{(p)}|_{x=-\mathcal{L}_p} = \frac{eD_n n_p}{L_n} (e^{eU/kT} - 1). \quad (9.152)$$

Следовательно, ВАХ тонкого p - n перехода описывается уравнением вида

$$J = e \left(\frac{D_n n_p}{L_n} + \frac{D_p p_n}{L_p} \right) (e^{eU/kT} - 1) = J_s (e^{eU/kT} - 1), \quad (9.153)$$

где плотность тока насыщения

$$\begin{aligned} J_s &= J_{sn} + J_{sp} = \frac{eD_n n_p}{L_n} + \frac{eD_p p_n}{L_p} = en_i^2 \left(\frac{D_n}{L_n p_p} + \frac{D_p}{L_p n_n} \right) = \\ &= e \left(\frac{n_p L_n}{\tau_n} + \frac{p_n L_p}{\tau_p} \right). \end{aligned} \quad (9.154)$$

Кривые распределения концентрации равновесных и избыточных носителей заряда и токов через p - n переход, построенные в соответствии с (9.144)—(9.152) для прямого и обратного смещений, изображены на рис. 9.16.

Из (9.153) следует, что при прямом смещении ток, текущий через p - n переход, экспоненциально возрастает с ростом приложенной

разности потенциалов, а обратный ток растет медленно и достигает тока насыщения. Таким образом, p - n переход обладает сильным выпрямляющим действием, которое тем лучше, чем меньше ток насыщения. Согласно (9.154) плотность тока насыщения J_s уменьшается с ростом концентрации основных носителей заряда n_n и p_p (т. е. с увеличением степени легирования n - и p -областей) и увеличением времени жизни неосновных носителей заряда τ_n и τ_p . Повышение температуры приводит к росту собственной концентрации n_i , а значит, и плотности тока насыщения J_s .

Как следует из рис. 9.15, при достижении критического значения отрицательного напряжения ток насыщения быстро увеличивается. Такой значительный рост тока обусловлен тем, что электроны и дырки в обедненном слое приобретают кинетическую энергию, достаточную для ударной ионизации валентных электронов. Возникающие при этом свободные носители заряда в свою очередь ускоряются полем и принимают участие в генерации электронно-дырочных пар. Происходит лавинное нарастание концентрации свободных носителей заряда, благодаря чему этот тип пробоя p - n перехода назван лавинным.

При приложении к p - n переходу внешнего напряжения состояние полупроводника будет неравновесным, характеризующимся квазиуровнями Ферми F_n и F_p . Если внешняя разность потенциалов U и не слишком велика, то избыточные концентрации носителей заряда будут существовать справа и слева от перехода лишь на расстоянии нескольких диффузионных длин L_n и L_p и в этих областях квазиуровни Ферми будут зависеть от координаты x (рис. 9.17).

Чтобы найти изменение положения квазиуровней Ферми, определим неравновесную концентрацию электронов в точке $x = -\mathcal{L}_p$. Она будет равна:

$$n(-\mathcal{L}_p) = n_p e^{eU/kT} = N_c e^{-[E_c - F_n(-\mathcal{L}_p)]/kT}, \quad (9.155)$$

где $F_n(-\mathcal{L}_p)$ — квазиуровень Ферми для электронов при $x = -\mathcal{L}_p$.

На расстоянии трех диффузионных длин ($3L_n$) от контакта концентрация равновесных электронов равна n_p и определяется соотношением вида

$$n(3L_n) = n_p = N_c e^{-(E_c - E_p)/kT}. \quad (9.156)$$

Подставляя значение n_p из (9.156) в (9.155), находим:

$$F_p = F_n(-\mathcal{L}_p) - eU. \quad (9.157)$$

Таким образом, при прямом смещении p - n перехода квазиуровень Ферми для электронов $F_n(x)$ в p -области на расстоянии $(2 \div 3) L_n$ понижается на величину eU , следовательно, удаляется от дна зоны проводимости E_c , а квазиуровень Ферми $F_p(x)$ в n -области на расстоянии $(2 \div 3) L_p$ повышается на eU , удаляясь от потолка валентной зоны E_v (рис. 9.17, а).

Следовательно, при прямом смещении p - n перехода дырочный квазиуровень Ферми в n -области приближается к верхнему краю валентной зоны, а квазиуровень Ферми для электронов в p -области

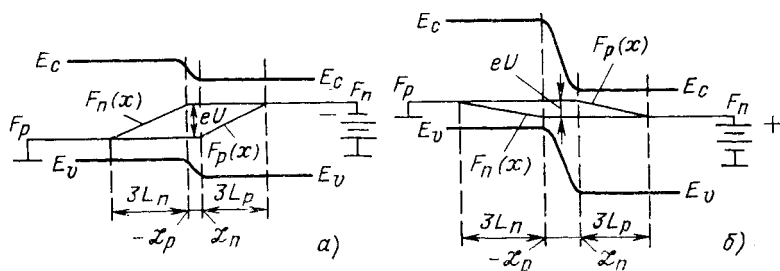


Рис. 9.17. Квазиуровень Ферми в p - n переходе при прямом (а) и обратном (б) смещении

поднимается ко дну зоны проводимости. Такое изменение квазиуровней Ферми и определяет возрастание концентраций неосновных носителей заряда у контакта.

При обратном смещении p - n перехода имеет место уменьшение количества неосновных носителей заряда вблизи контакта и квазиуровни Ферми $F_n(x)$ и $F_p(x)$ в приконтактных областях удаляются от соответствующих зон (рис. 9.17, б).

Рассмотрим теперь соотношение токов, текущих через p - n переход. На основании (9.151) и (9.152) отношение электронной и дырочной составляющих плотностей токов через p - n переход

$$\frac{J_n}{J_p} = \frac{J_{sn}}{J_{sp}} = \frac{D_n n_p L_n}{D_p p_n L_p} = \frac{\mu_n n_p L_p}{\mu_p p_n L_n} = \frac{\sigma_n L_p}{\sigma_p L_n}. \quad (9.158)$$

Из этого равенства следует, что отношение электронного и дырочного токов, текущих через p - n переход, определяется главным образом отношением концентраций основных носителей заряда в n - и p -областях. Если степень легирования обеих частей полупроводника приблизительно одинакова ($\sigma_n \approx \sigma_p$), то такой p - n переход инжектирует в равной степени электроны и дырки и плотность общего тока определяется выражением

$$J = (J_{sn} + J_{sp}) (e^{eU/kT} - 1). \quad (9.159)$$

В том случае, когда p -область легирована значительно сильнее n -области ($\sigma_p \gg \sigma_n$), то $J_{sp} \gg J_{sn}$ и через p - n переход течет в основном дырочный ток. При $\sigma_n \gg \sigma_p$, напротив, основной вклад в общий ток вносит электронная составляющая тока.

9.11. n^+ - n И p^+ - p ПЕРЕХОДЫ

Большое практическое значение имеют переходы, образованные между двумя областями полупроводника с одинаковым типом электропроводности, но обладающими различными значениями удельной проводимости. Могут быть созданы электронно-электронные переходы, или n^+ - n переходы, в которых n^+ -область содержит большее количество легирующей примеси — доноров, а также дырочно-

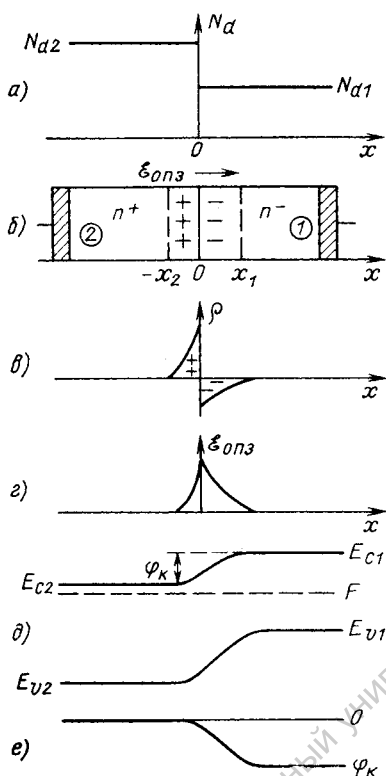


Рис. 9.18. Распределение легирующей примеси (а), структура n^+-n перехода (б), плотности объемного заряда (в), распределение напряженности электрического поля (г), энергетические зоны (д) и электростатический потенциал (е) в n^+-n переходе в состоянии термодинамического равновесия

дырочные переходы, или p^+-p переходы, в которых одна часть перехода легирована акцепторами значительно сильнее, чем другая. n^+-n и p^+-p переходы, которые получили название изотипных переходов, практически не обладают выпрямляющими свойствами.

Рассмотрим идеализированный резкий n^+-n переход, концентрация доноров в котором меняется скачком от значения N_{d2} в n -области до меньшего значения N_{d1} в n^+ -области (рис. 9.18, а). В результате диффузии электронов приконтактная n^+ -область обедняется электронами и положительный объемный заряд в ней образован нескомпенсированными ионами донорной примеси, а прикон-

тактная n -область обогащается электронами, создающими отрицательный объемный заряд. Границы области пространственного заряда (ОПЗ) на рис. 9.18, б изображены штриховыми линиями. Разделение зарядов и образование ОПЗ (рис. 9.18, в) порождает электрическое поле (рис. 9.18, г), что приводит к искривлению энергетических зон (рис. 9.18, д), а следовательно, к возникновению потенциального барьера (рис. 9.18, е).

В состоянии термодинамического равновесия положение уровня Ферми F постоянно по всему образцу полупроводника, диффузионный ток через n^+-n переход равен нулю и в отсутствие вырождения в любой точке кристалла $n_0 p_0 = n_i^2$. Для того чтобы определить контактную разность потенциалов φ_K , предположим, что n^+ - и n -области имеют значительную толщину и донорная примесь в них полностью ионизована. В этом случае равновесные концентрации электронов в глубине n^+ - и n -областей согласно (4.66) будут равны:

$$n_{20} = N_c e^{-(E_{c2}-F)/kT} = N_{d2}; \quad (9.160)$$

$$n_{10} = N_c e^{-(E_{c1}-F)/kT} = N_{d1}, \quad (9.161)$$

где индексы 2 и 1 относятся соответственно к n^+ - и n -областям, а изменение электростатического потенциала на основе (9.103) соответствует

$$\varphi_1 - \varphi_2 = -\frac{E_{c1} - E_{c2}}{e} = -\frac{kT}{e} \ln \left(\frac{n_{20}}{n_{10}} \right) = -\frac{kT}{e} \ln \left(\frac{N_{d2}}{N_{d1}} \right). \quad (9.162)$$

Для n^+ - n перехода существенным является то, что область пространственного заряда, в которой сосредоточено электрическое поле, простирается в основном в слабо легированную n -область. Она обогащена электронами, и ее толщина соответствует дебаевской длине экранирования (8.35).

Если положить $\varphi_2 = 0$ (рис. 9.18, e), то для φ_K получим:

$$\varphi_K = -\frac{kT}{e} \ln \left(\frac{N_{d2}}{N_{d1}} \right). \quad (9.163)$$

При наложении внешнего напряжения на n^+ - n переход основная часть напряжения падает на высокоомной n -области полупроводника. Тем не менее небольшая часть внешнего напряжения падает на области пространственного заряда, увеличивая или уменьшая высоту потенциального барьера. Когда на переходе положительное напряжение, т. е. n -область заряжена положительно, то в этом случае имеет место понижение высоты потенциального барьера, поскольку напряженность внешнего электрического поля $\mathcal{E}_1$ в n -области противоположна по направлению напряженности поля пространственного заряда $\mathcal{E}_{опз}$. При обратном смещении n^+ - n перехода, когда n -область заряжена отрицательно и $\mathcal{E}_1$ по направлению совпадает с $\mathcal{E}_{опз}$, высота потенциального барьера повышается.

Рассмотрим физические процессы, приводящие к нелинейности ВАХ n^+ - n перехода. Будем считать, что внешние контакты и структура являются омическими, на которых концентрации дырок и электронов равны равновесным значениям. Кроме того, положим, что протяженность n^+ - и n -областей значительно больше дебаевской и диффузионной длин, так что внешние омические контакты практически не влияют на распределение электронов и дырок вблизи n^+ - n перехода.

Поскольку концентрация электронов в n^+ -области n_{20} очень велика, напряженность поля $\mathcal{E}_2$ в глубине этой области ничтожно мала по сравнению с $\mathcal{E}_1$ в глубине n -области при любой полярности внешнего напряжения. Так как $p_{20} \ll p_{10}$, то плотность дрейфового тока дырок $J_{p_2 \text{ др}} \ll J_{p_1 \text{ др}}$.

При обратном смещении n^+ - n перехода напряженность электрического поля $\mathcal{E}_1$, которая совпадает по направлению с напряженностью поля области пространственного заряда (рис. 9.18, e), быстро уносит дырки от границы x_1 в глубь n -области, их концентрация не может быть возмещена тепловой генерацией и приходом очень небольшого количества дырок из n^+ -области. В силу этого n -область обедняется дырками (кривая 1 на рис. 9.19, a). Сохранение электронейтральности образца будет в том случае, если количество электронов в n -области уменьшится на такое же значение, причем максималь-

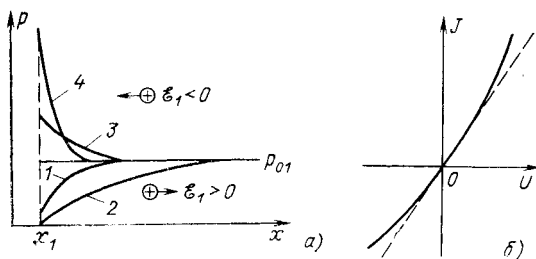


Рис. 9.19. Распределение дырок в слаболегированной n -области (а) и вольт-амперная характеристика (б) n^+-n перехода

ное обеднение ими будет у границы x_1 , т. е. будет иметь место явление экстракции носителей заряда. При этом концентрация дырок в n -области уменьшается по экспоненциальному закону с постоянной спада L_1 , определяемой формулой (8.59). С учетом (8.28), (8.44) и (8.45) можно записать

$$L_1 = \frac{2L_{p_1}^2}{\sqrt{L_{\mathcal{E}_1}^2 + 4L_{p_1}^2 - L_{\mathcal{E}_1}}} = L_{p_1} \left[\sqrt{1 + \left(\frac{L_{p_1} \mathcal{E}_1}{2\varphi_T} \right)^2} + \frac{L_{p_1} \mathcal{E}_1}{2\varphi_T} \right], \quad (9.164)$$

где $\varphi_T = kT/e$ — температурный потенциал.

С ростом напряженности электрического поля область обеднения носителями заряда будет распространяться дальше в n -область образца (кривая 2 на рис. 9.19, а). Это приводит к некоторому возрастанию сопротивления n -области и отклонению обратной ветви ВАХ n^+-n перехода от линейной зависимости, как показано на рис. 9.19, б, которое будет наблюдаться в области больших токов. Влияние процесса экстракции носителей заряда на увеличение сопротивления n -области при обратном смещении n^+-n перехода ощутимо только тогда, когда концентрация дырок в n -области сравнима с концентрацией электронов, т. е. когда проводимость n -области близка к собственной. Следовательно, n^+-n переходы можно использовать в качестве контактов, когда необходимо избежать инжекции дырок в полупроводник n -типа.

При прямом направлении смещения n^+-n перехода напряженность поля $\mathcal{E}_1$ направлена противоположно напряженности поля $\mathcal{E}_{опз}$ и поэтому дырки в n -области движутся по направлению к потенциальному барьеру. При небольших прямых смещениях, когда $U < \varphi_K$, высота потенциального барьера уменьшается незначительно и дырки практически не проникают через барьер в n^+ -область. По этой причине прилегающая к границе часть n -области полупроводника обогащается дырками, как показано на рис. 9.19, а (кривые 3 и 4), т. е. имеет место аккумуляция дырок, концентрация которых в зависимости от координаты будет уменьшаться по экспоненциальному закону с постоянной спада L_2 (8.62), соответствующей диффузии дырок против поля, равной

$$L_2 = \frac{2L_{p_1}^2}{\sqrt{L_{\mathcal{E}_1}^2 + 4L_{p_1}^2 + L_{\mathcal{E}_1}}} = L_{p_1} \left[\sqrt{1 + \left(\frac{L_{p_1} |\mathcal{E}_1|}{2\varphi_T} \right)^2} - \frac{L_{p_1} |\mathcal{E}_1|}{2\varphi_T} \right]. \quad (9.165)$$

Электронейтральность образца будет наблюдаться при увеличении концентрации основных носителей заряда — электронов на $\Delta n = \Delta p$. При этом происходит изменение проводимости n -области и сопротивление n^+-n перехода немного уменьшается, а поэтому прямая ветвь ВАХ становится нелинейной (рис. 9.19, б). Эта нелинейность тем больше, чем больше прямое смещение, приводящее к значительному снижению высоты потенциального барьера, что обусловит увеличение количества электронов, проникающих из n^+ - в n -область, а следовательно, вызовет дополнительное снижение сопротивления n^+-n перехода.

9.12. ГЕТЕРОПЕРЕХОДЫ

Гетеропереходом называют переход, образующийся на границе контакта двух полупроводников с различной шириной запрещенной зоны. Существуют два типа гетеропереходов: переход между полупроводниками p - и n -типа и переход между двумя материалами n -типа или между двумя полупроводниками p -типа. Подбор таких материалов представляет сложную задачу, ибо необходимо выполнение многих условий «совместимости», важнейшими из которых являются наличие одинаковых постоянных решеток и одинаковых коэффициентов расширения контактирующих веществ. Гетеропереходы могут быть образованы совершенно различными материалами либо один из материалов представляет собой полупроводниковое соединение, а другой — его сплав с родственным соединением, величины запрещенных зон которых существенно отличаются друг от друга. В табл. 9.1 представлены параметры полупроводников для шести гетеропереходов.

Рассмотрим идеальный p - n переход, у которого материал p -типа имеет более широкую запрещенную зону, чем полупроводник n -

Таблица 9.1. Параметры полупроводников, образующих гетеропереходы

Гетеро- переход	Полу- провод- ник	E_g , эВ, при $T = 300$ К	Пара- метр ре- шетки, им	Кэффи- циент линейного расширения при 300 К, 10^{-6} К $^{-1}$	Энергия электронно- го сродства, эВ	Относитель- ная диэлек- трическая постоянная ϵ_r
GaAs—Ge	GaAs	1,43	0,5653	5,8	4,07	11,5
	Ge	0,67	0,5658	5,7	4,13	16,0
AlSb—GaSb	AlSb	1,6	0,6136	3,7	3,65	10,3
	GaSb	0,68	0,6095	6,9	5,06	14,8
AlAs—GaAs	AlAs	2,15	0,5661	5,2	—	—
	GaAs	1,43	0,5653	5,8	4,07	11,5
GaP—Si	GaP	2,25	0,5451	5,3	4,3	3,4
	Si	1,11	0,5431	2,33	4,01	12
ZnSe—Ge	ZnSe	2,67	0,5667	7,0	4,09	9,1
	Ge	0,66	0,5658	5,7	4,13	16,0
ZnSe—GaAs	ZnSe	2,67	0,5669	7,0	4,09	9,1
	GaAs	1,43	0,5653	5,8	4,07	11,5

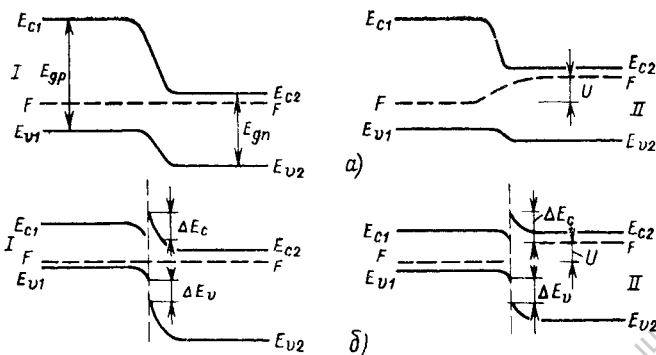


Рис. 9.20. Диаграммы энергетических зон идеального (а) и резкого (б) p - n гетеропереходов при термодинамическом равновесии (I) и положительном смещении (II)

типа, т. е. $E_{gp} > E_{gn}$. Структура такого p - n гетероперехода в случае термодинамического равновесия, когда энергия уровня Ферми F одинакова по обе стороны перехода, показана на рис. 9.20, а, структура I.

Предположим, что к идеальному p - n гетеропереходу приложено внешнее напряжение U в прямом направлении (плюс источника подключен к p -области). В этом случае, как следует из рис. 9.20, а, структура II, потенциальный барьер, который должен преодолеть дырки при переходе из p -области в n -область, заметно понижается. И если приложенное напряжение U достаточно велико, барьер может быть понижен до нуля. Следовательно, дырочный ток в прямом направлении сильно возрастает. Но потенциальный барьер, который должны преодолеть электроны при переходе из n -области в p -область, остается достаточно высоким, а поэтому электронный ток в прямом направлении будет довольно мал. ВАХ такого p - n гетероперехода в прямом направлении описывается уравнением (9.153), но значение плотности тока насыщения отличается от J_s для p - n перехода (9.154). В обратном направлении ток не насыщается, как в обычном p - n переходе, а растет с увеличением $-U$, оставаясь много меньше «прямого» тока.

В том случае, если два полупроводника имеют различную ширину запрещенной зоны E_g , различную относительную диэлектрическую проницаемость ϵ_r , различную работу выхода Φ и различное электронное сродство χ , то равновесная диаграмма такого резкого p - n перехода будет такая, как изображено на рис. 9.20, б, структура I. Электрическое поле, обуславливающее наклон зон на границе раздела, терпит разрыв вследствие различия в величинах ϵ_r , это обусловит и разрыв краев энергетических зон на границе раздела. В этом случае барьер для электронов значительно меньше, чем для дырок, поэтому доминирующими носителями будут электроны.

Гетеропереходы p - n находят широкое применение в качестве эффективных инжекторов дырок в материал n -типа. Это устраняет

необходимость в сильно легированной области p -типа. Они используются при создании полупроводниковых лазеров. Область p -типа, имеющая более широкую запрещенную зону, прозрачна для рекомбинационного излучения из области p -типа. Кроме того, в области p -типа отсутствует поглощение света свободными носителями заряда, поскольку она не сильно легирована.

В случае различия в значениях постоянной решетки на границе двух полупроводников гетероперехода будет существовать механическое напряжение. Это приведет к появлению в переходном слое дефектов, которые будут являться ловушками как для дырок, так и для электронов. Эти ловушки удерживают уровень Ферми в переходном слое посередине запрещенной зоны обоих полупроводников (см. рис. 4.22, б) и вызывают подъем зон со стороны n -типа и их понижение со стороны p -типа, так же как в случае контакта металл—полупроводник (см. рис. 9.7 и 9.8). В результате вид энергетических зон p - n и n^+ - n гетеропереходов станет таким, как на рис. 9.21.

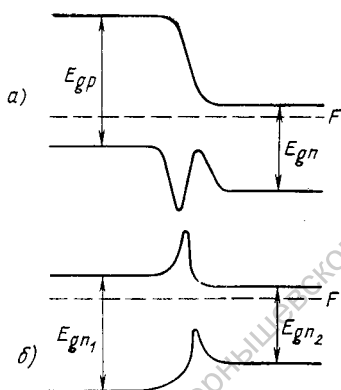


Рис. 9.21. Вид энергетических зон p - n (а) и n^+ - n гетероперехода (б) при наличии ловушек в переходной области

9.13. КОНТАКТ ВЫРОЖДЕННЫХ ЭЛЕКТРОННОГО И ДЫРОЧНОГО ПОЛУПРОВОДНИКОВ. ТУННЕЛЬНЫЙ ДИОД

В p - n переходе, образованном двумя областями невырожденного электронного и дырочного полупроводников, ток обуславливается прохождением электронов над потенциальным барьером. В случае вырожденного полупроводника с концентрацией примесей в обеих областях порядка 10^{18} — 10^{20} см $^{-3}$ переходный слой будет очень тонкий и возможно прохождение электронов через p - n переход в результате туннельного эффекта, а поэтому вид ВАХ будет принципиально отличаться от обычной диодной характеристики (9.153). Благодаря возникновению туннельного эффекта наблюдается резкий рост тока при обратной полярности напряжения (к p -области присоединен отрицательный вывод источника), а на участке прямого смещения появляется область отрицательного дифференциального сопротивления (рис. 9.22). Значительный туннельный ток возникает в p - n переходах толщиной около 10 нм, когда при контактной разности потенциалов около 1 В напряженность поля в переходе приближается к 10^6 В/см (см. § 6.12). Такая толщина перехода для большинства полупроводников наблюдается при концентрации примеси, лежащей в указанном выше диапазоне.

Рассмотрим ход ВАХ туннельного диода на различных участках. На рис. 9.23, а приведена упрощенная энергетическая диаграмма контакта вырожденных полупроводников при отсутствии внешнего

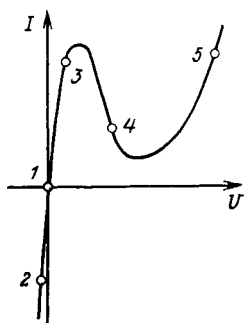


Рис. 9.22. Вольт-амперная характеристика туннельного диода

смещения. Штриховкой показаны состояния, занятые электронами. На рис. 9.23, а и на последующих рисунках представлены также плотности свободных p и заполненных n (заштрихованы) состояний в валентной зоне и зоне проводимости, индексы которых s и v относятся соответствующим зонам. Для простоты принято, что плотность квантовых состояний $N(E)$ изменяется по степенному закону для валентной зоны $N_p(E) \sim (E_v - E)^{1/2}$ (4.11) и для зоны проводимости $N_n(E) \sim (E - E_c)^{1/2}$ (4.9) (эти зависимости изображены пунктиром). Следует отметить, что это предположение несправедливо для вырожденного полупроводника, однако оно позволяет значительно упростить качественное рассмотрение характеристики при-

бора, изображенной на рис. 9.22. Значения ξ_n и ξ_p характеризуют степень вырождения соответствующих областей полупроводника и определяют суммарное перекрытие разрешенных энергетических зон. Благодаря наличию такого перекрытия электроны могут переходить из одной области в другую за счет туннелирования сквозь потенциальный барьер, форма которого близка к треугольной. Вероятность туннелирования D определяется выражением (6.221). Она слабо зависит от формы потенциального барьера, ее можно также считать не зависящей от направления движения электронов. При отсутствии внешнего смещения суммарный ток через p - n переход равен нулю, что соответствует точке 1 на ВАХ рис. 9.22.

Допустим, что теперь к p - n переходу приложено обратное смещение. В этом случае, как следует из рис. 9.24, все энергии в n -области снижаются относительно p -области и поток электронов из p -области в n -область резко возрастает, так как увеличивается количество заполненных уровней в полупроводнике p -типа, против которых при той же энергии лежат свободные уровни в зоне проводимости материала n -типа. Поток электронов в обратном направлении при этом уменьшится. Суммарный ток электронов обозначен стрелкой на рис. 9.24, а, этому режиму соответствует точка 2 на рис. 9.22. Уве-

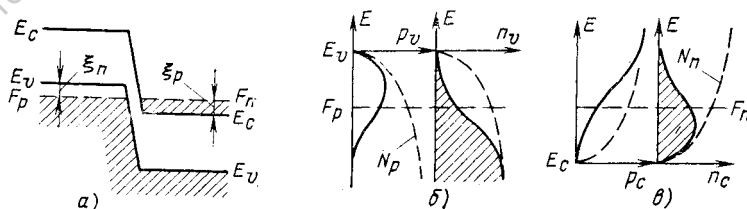


Рис. 9.23. Туннельный диод без внешнего смещения:

зонная диаграмма (а), плотности свободных и заполненных состояний в валентной зоне (б) и в зоне проводимости (в)

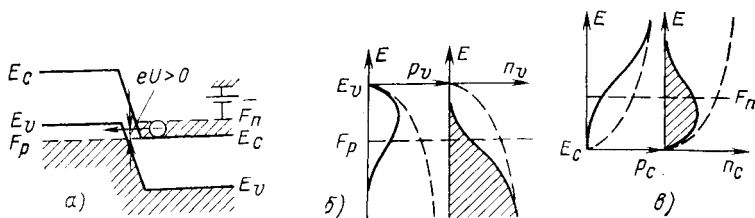


Рис. 9.24. Туннельный диод при обратном смещении:

зонная диаграмма (а), плотности свободных и заполненных состояний в валентной зоне (б) и в зоне проводимости (в)

личение обратного напряжения сопровождается ростом туннельного тока.

При небольших положительных напряжениях, как видно из рис. 9.25, возрастает количество электронов, туннелирующих из n -области в p -область (эти переходы указаны стрелкой) при одновременном снижении встречного потока. В этом нетрудно убедиться, сравнивая концентрации заполненных и свободных энергетических состояний по обеим сторонам перехода. Описанному режиму соответствует точка 3 на рис. 9.22.

При дальнейшем повышении прямого напряжения перекрытие разрешенных зон проходит через максимум и начинает уменьшаться, что ведет за собой снижение туннельного тока (рис. 9.26). При этом рабочая точка прибора на ВАХ (см. рис. 9.22) будет перемещаться к точке 4. Увеличение входного напряжения до величины $U = \frac{1}{e}(\xi_n + \xi_p)$, когда E_c совпадает с E_v , приведет к уменьшению туннельного тока до нуля.

Однако по мере повышения прямого смещения на p - n переходе высота потенциального барьера понижается и будет возрастать диффузионный ток основных носителей заряда, способных преодолеть снижающийся потенциальный барьер. Ток будет увеличиваться по тому же закону, что и в обычном диоде (точка 5 на рис. 9.22).

Из рис. 9.27 следует, что наибольшее различие в ВАХ, рассчитанной теоретически и измеренной экспериментально, проявляется в области минимума тока, где его реальная величина существенно

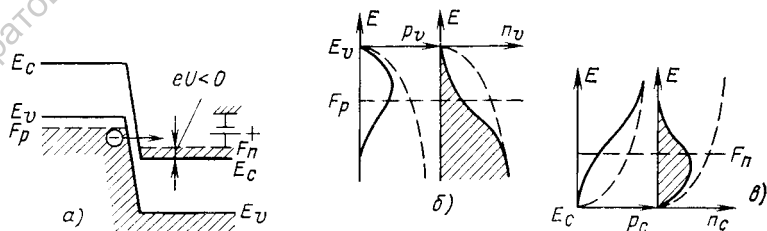


Рис. 9.25. Туннельный диод при малом прямом смещении:

зонная диаграмма (а), плотности свободных и заполненных состояний в валентной зоне (б) и в зоне проводимости (в)

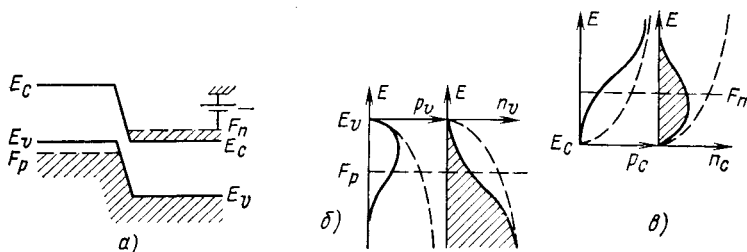


Рис. 9.26. Туннельный диод при прямом смещении, соответствующем спаду тока: зонная диаграмма (а), плотности свободных и заполненных состояний в валентной зоне (б) и в зоне проводимости (в)

превосходит сумму расчетных значений для туннельной и диффузионной составляющих. В этой области существенное влияние оказывает изменение плотности состояний вблизи краев запрещенной зоны вырожденного полупроводника. Этот эффект проявляется в возникновении хвостов плотности состояний, заходящих в запрещенную зону (см. § 4.9). Учет данного явления показывает, что туннельный ток не спадает до нуля при $U = \frac{1}{e} (\xi_n + \xi_p)$, так как при этом еще наблюдается перекрытие разрешенных состояний в n - и p -областях.

Вторым существенным фактором, объясняющим увеличение реального тока в области минимума ВАХ туннельного диода, является наличие в запрещенной зоне полупроводника глубоких примесных уровней, которые обеспечивают избыточный туннельный ток при прямом смещении. На рис. 9.28 A — один из возможных путей туннелирования электронов на примесный уровень в p -области с дальнейшим их переходом на свободные уровни валентной зоны; B — переход с примесного уровня в запрещенной зоне n -области в валентную зону материала p -типа. Освободившееся примесное состояние занимает электрон из зоны проводимости. Протекающий туннельный ток зависит от суммарной концентрации примесных состояний и их расположения в запрещенной зоне полупроводника.

Благодаря наличию участка отрицательного дифференциального сопротивления туннельный диод может быть использован для уси-

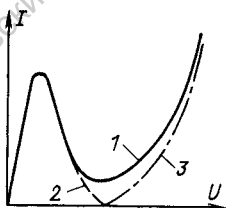


Рис. 9.27. Прямая ветвь вольт-амперной характеристики туннельного диода:

1 — экспериментальная кривая; 2 — туннельный ток; 3 — диффузионный ток

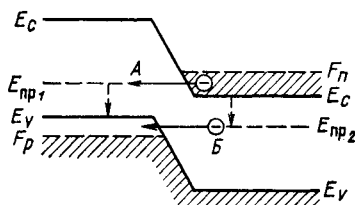


Рис. 9.28. Туннелирование электронов через примесные уровни

ления и генерации колебаний. Поскольку прибор работает на основных носителях заряда, его инерционность ограничивается временем диэлектрической релаксации (см. § 7.3), которое для германия имеет порядок 10^{-13} с. Практически быстрое действие диода оказывается значительно ниже из-за влияния емкости структуры и других элементов схемы. Тем не менее туннельные диоды способны работать в схемах генераторов и усилителей СВЧ-диапазона. Кроме того, они используются в качестве быстродействующих переключателей с двумя устойчивыми состояниями. Принципиальным недостатком этих приборов, ограничивающим их практическое применение, является малая выходная мощность, что связано с незначительным изменением тока и напряжения на участке отрицательного дифференциального сопротивления. Наличие избыточного тока отрицательно сказывается на параметрах диода, так как приводит к снижению перепада тока на рабочем участке ВАХ.

9.14. ОМИЧЕСКИЙ ПЕРЕХОД

При измерении электрофизических параметров полупроводника исследуемый образец должен иметь контакты для подведения тока, не влияющие на поведение носителей заряда в объеме полупроводника. Механические, прижимные контакты для этой цели не годятся, так как наличие слоя окисла и загрязнений на поверхности образца приводит, как правило, к повышенному сопротивлению контакта, а измеряемые характеристики являются невоспроизводимыми. Поэтому необходим физический контакт, электрическое сопротивление которого мало и не зависит от направления тока в заданном рабочем диапазоне значений токов. Такие электрические контакты называют омическими переходами. Дополнительными условиями таких контактов является стабильность их электрических и механических свойств и отсутствие инжекции неосновных носителей заряда. Хорошие омические переходы должны иметь удельное сопротивление около или меньше 10^{-5} Ом·см².

В качестве омического перехода может быть использован переход металл—полупроводник с антизапорным слоем, при этом у металла термоэлектронная работа выхода должна быть меньше, чем у полупроводника n -типа (см. рис. 9.7, б), или больше, чем у полупроводника p -типа (рис. 9.8, а). Однако комбинаций металл—полупроводник с такими свойствами очень мало. Кроме того, такой переход обладает заметной инжекцией неосновных носителей заряда, чему способствует повышенная генерация их через дефекты на поверхности полупроводника и направление приконтактного электрического поля, выталкивающего эти носители заряда в объем полупроводника.

Большинство омических переходов создается на основе n^+-n или p^+-p переходов. Концентрация легирующей примеси в сильно легированном слое должна быть настолько высока, чтобы между металлом и n^+ -полупроводником создавалась обедненная область такой толщины, при которой носители заряда благодаря туннельному эф-

фекту беспрепятственно переходили через потенциальный барьер, и в силу этого сопротивление перехода будет малым. Из-за низкой концентрации дырок в вырожденном n^+ -слое инжекция их в слабо легированную n -область практически отсутствует.

Образованию омического перехода способствует высокая скорость рекомбинации носителей заряда на границе полупроводника и металла, что достигается нарушением поверхности полупроводника (например, шлифовкой) перед нанесением на нее металла. В этом случае вблизи поверхности полупроводника создаются дефекты кристаллической решетки, действующие как эффективные центры рекомбинации. При этом концентрация основных и неосновных носителей заряда вблизи поверхности полупроводника будет примерно равна равновесной.

Создание качественных омических переходов, несмотря на относительную простоту рецептов их изготовления, представляет собой сложную технологическую задачу. Ряд металлов, пригодных по своим электрофизическим свойствам для изготовления омических контактов, не смачивают поверхность полупроводника, скатываясь в шарики при нагреве, обладают плохой адгезией, химически активны и вступают в реакцию как с полупроводником, так и с кислородом воздуха, окисляясь при этом. Поэтому изготовление хороших омических переходов представляет собой достаточно сложную химико-технологическую проблему.

ГЛАВА ДЕСЯТАЯ

ПОВЕРХНОСТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

10.1. ПРИРОДА ПОВЕРХНОСТНЫХ УРОВНЕЙ

В гл. 2 при рассмотрении энергетического спектра электрона в твердом теле в приближении сильносвязанных электронов предполагалось, что везде в кристалле сохраняется строгая периодичность кристаллического потенциала, а ограниченность объема кристалла учитывалась введением циклических граничных условий Борна—Кармана. Последние приводили к дискретности энергетического спектра в пределах разрешенных зон энергий. Однако более детальное рассмотрение показывает, что влияние конечных размеров кристалла наличием границ не исчерпывается. Действительно, поверхность есть естественное нарушение периодичности потенциала и, следовательно, можно ожидать появления особенностей в спектре разрешенных состояний электрона кристалла. Таммом было показано, что обрыв периодичности кристаллического потенциала на поверхности приводит к появлению локализованных состояний, энергетические уровни которых располагаются в запрещенной зоне. Эти состояния называют **поверхностными состояниями** или **уровнями Тамма**. Поскольку обрыв потенциала происходит в каждой цепочке атомов, нормальной к поверхности, то, оче-

видно, плотность таммовских уровней равна плотности поверхностных атомов, т. е. 10^{15} см^{-2} .

К такой же плотности состояний приводят поверхностные уровни, теоретически предсказанные Шокли и интерпретирующиеся как насыщенные валентности поверхностных атомов кристалла.

Оба типа состояний относятся к идеализированной модели поверхности и могут рассматриваться лишь применительно к так называемой атомарно-чистой поверхности. Однако даже в случае атомарно-чистой поверхности структура поверхности существенно отличается от идеализированной модели из-за перегруппировки поверхностных атомов, вызванной взаимным насыщением свободных валентных связей поверхностных атомов кристалла.

Реальная поверхность полупроводника, с которой приходится иметь дело, весьма далека от идеальной, ибо на ней практически всегда имеют место различного рода макроскопические и микроскопические структурные дефекты, связанные с условиями обработки (резка, полировка, шлифовка и т. д.) и роста кристалла (огранка кристалла и другого рода особенности рельефа поверхности макроскопических и микроскопических размеров). Кроме того, реальная поверхность полупроводника находится в постоянном контакте с окружающей средой, различного рода химическими соединениями, применяемыми в качестве травителей, в результате чего на поверхности возможна адсорбция посторонних атомов и молекул из этих источников, появление окисных пленок как результат окислительно-растворительных реакций при травлении и т. д. Все это приводит к появлению локализованных на поверхности полупроводника состояний, которые в зависимости от степени сродства к электрону и дырке, положения уровня Ферми на поверхности могут проявлять себя как донорные или акцепторные центры захвата или рекомбинационные ловушки электронно-дырочных пар. При наличии, например, донорных состояний на поверхности, которые, как известно, могут быть либо нейтральными, либо заряженными положительно при отдаче электрона в зону проводимости, поверхность полупроводника будет заряжена положительно. При наличии акцепторных состояний поверхность полупроводника будет заряжена отрицательно, поскольку акцепторные состояния могут быть либо нейтральны (когда пусты), либо отрицательно заряжены (когда заполнены электронами). Величина заряда Q_{ss} на поверхности зависит от концентрации поверхностных состояний N_s и функций распределения для ловушек захвата, которые определяются значением электрохимического потенциала на поверхности полупроводника по отношению к энергетическому положению уровня ловушки захвата.

В условиях термодинамического равновесия полупроводник в целом электронейтрален. В приповерхностной области в присутствии заряда в поверхностных состояниях электронейтральность обеспечивается тем, что электрическое поле вблизи поверхности, вызванное зарядом Q_{ss} , приводит к перераспределению подвижных носителей заряда в приповерхностной области полупроводника, в результате чего в ней возникает пространственный заряд Q_{sp} , равный по

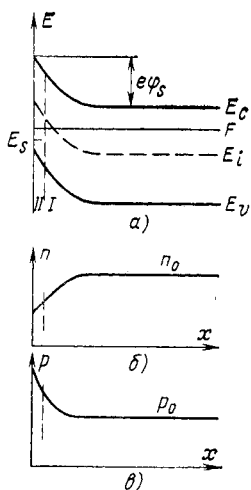


Рис. 10.1. Энергетическая диаграмма (а) и распределение концентрации электронов (б) и дырок (в) в приповерхностной области полупроводника n -типа:

I — область обеднения;
 II — область инверсии

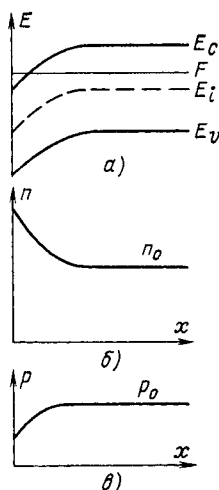


Рис. 10.2. Энергетическая диаграмма (а) и распределение концентрации электронов (б) и дырок (в) в приповерхностной области полупроводника в случае обогащения

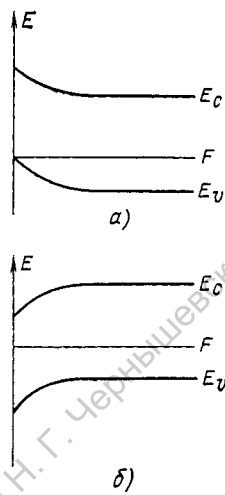


Рис. 10.3. Энергетическая диаграмма приповерхностной области полупроводника p -типа в случае обогащения (а) и обеднения (б)

величине и противоположный по знаку заряду в поверхностных состояниях. Таким образом, возникает двойной электрический слой, экранирующий объем полупроводника от действия поля.

В металле, где концентрация свободных электронов составляет 10^{22} см^{-3} , нейтрализация поверхностного заряда происходит на расстоянии $10^{-8}—10^{-7} \text{ см}$ и такой тонкий слой пространственного заряда не может существенно сказываться на свойствах всего кристалла. У полупроводников концентрация свободных носителей заряда значительно меньше, поэтому область пространственного заряда достаточно обширна. Например, для германия, у которого концентрация поверхностных уровней составляет 10^{11} см^{-2} , при концентрации носителей заряда 10^{15} см^{-3} нейтрализация поверхностного заряда происходит на расстоянии 10^{-4} см , а в собственном германии ($n_0 \approx 10^{13} \text{ см}^{-3}$) — уже на расстоянии $0,1 \text{ мм}$. Наличие у полупроводников поверхностного заряда изменяет его энергетическую схему в приповерхностной области.

Рассмотрим образование приповерхностного слоя объемного заряда на примере электронного полупроводника, на поверхности которого имеются акцепторные уровни E_a , как это показано на рис. 10.1. При заполнении акцепторных поверхностных состояний электронами на поверхности полупроводника возникает отрицательный заряд, а в его приповерхностном слое при этом появится обед-

ненный электронами слой, обладающий положительным пространственным зарядом. Наличие двойного электрического слоя приводит к появлению электрического поля, направленного к поверхности полупроводника, которое вызывает в приповерхностном слое полупроводника изгиб его энергетических зон вверх. Если через $e\varphi_s$ обозначить значение изгиба зоны проводимости в поверхностной области, то φ_s — электростатический поверхностный потенциал. В приповерхностном слое такого полупроводника в зависимости от положения границы его энергетических зон относительно уровня Ферми в общей сложности может быть область обеднения, для которой проводимость меньше, чем в объеме полупроводника, и область, в которой имеет место изменение типа проводимости, так называемый *и-в-е-р-с-и-о-н-н-ы-й-с-л-о-й* (рис. 10.1, а, области I и II соответственно). Образование инверсионного слоя зависит от степени легирования образца.

В том случае, если у электронного полупроводника на поверхности имеются донорные состояния, то его энергетические зоны в приповерхностной области изгибаются вниз и образуется область обогащения. В приповерхностном слое такого полупроводника концентрация основных носителей заряда больше, чем в объеме (рис. 10.2).

У акцепторных полупроводников обеднение наблюдается в том случае, когда поверхностные состояния захватывают дырки и зоны изгибаются вниз, а при захвате электронов имеет место обогащение, при котором зоны изгибаются вверх (рис. 10.3).

10.2. ТЕОРИЯ СЛОЯ ПРОСТРАНСТВЕННОГО ЗАРЯДА

Поскольку в приповерхностной области полупроводника имеет место изгиб энергетических зон, то концентрация подвижных носителей заряда зависит от координаты в направлении, нормальном к поверхности (рис. 10.1, б, в). Для выяснения конкретного вида этих зависимостей необходимо решить уравнение Пуассона. В § 9.4 подобное решение проведено для случая обедненных слоев на контакте металл — полупроводник в пренебрежении вкладом неосновных носителей заряда в так называемое приближение обедненного слоя Шоттки, которое приводит к квадратичной зависимости потенциала от координаты (9.52). Это решение, однако, не описывает $\varphi(x)$ для случая обогащенных и инверсионных слоев. Поэтому необходимо решить уравнения Пуассона в общем виде, не пренебрегая вкладом в плотность объемного заряда подвижных носителей заряда.

Прежде чем решать уравнения Пуассона, введем обозначения, принятые при рассмотрении поверхностных явлений (рис. 10.4). Удобно вести отсчет всех

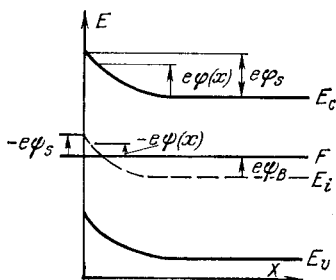


Рис. 10.4. Энергетическая диаграмма поверхностной области полупроводника *n*-типа

величин, характеризующих полупроводник в области пространственного заряда и вне ее, от F_i — уровня Ферми для собственного полупроводника [обычно F_i называют серединой запрещенной зоны, что, однако, полностью соответствует действительности лишь в случае равенства эффективных масс плотностей состояния $m_{\bar{n}}$ и $m_{\bar{p}}$ (см. соотношение (4.88)]. Потенциал Ψ определяется в общем случае:

$$e\Psi = F - F_i, \quad (10.1)$$

где F — уровень Ферми для данного полупроводника.

Электростатический потенциал в объеме полупроводника обозначим через Ψ_B , на его поверхности — через Ψ_s . Потенциал в произвольной точке пространственного заряда $\varphi(x) = \Psi(x) - \Psi_B$, на поверхности $\varphi_s = \Psi_s - \Psi_B$. Концентрация электронов и дырок в области пространственного заряда может быть выражена через φ и Ψ следующим образом:

$$n = N_c e^{-(E_c + U - F)/kT} = n_0 e^{e\varphi/kT} = n_i e^{e\Psi_B/kT} e^{e(\Psi - \Psi_B)/kT} = n_i e^{e\Psi/kT}. \quad (10.2)$$

Так как $U = -e\varphi$ (в соответствии с рис. 10.1 и 10.4 $\varphi < 0$), то

$$p = p_0 e^{-e\varphi/kT} = n_i e^{-e\Psi_B/kT} e^{e(\Psi_B - \Psi)/kT} = n_i e^{-e\Psi/kT}. \quad (10.3)$$

Поверхностные концентрации электронов и дырок

$$n_s = n_0 e^{e\varphi_s/kT} = n_i e^{e\Psi_s/kT}, \quad (10.4)$$

$$p_s = p_0 e^{-e\varphi_s/kT} = n_i e^{-e\Psi_s/kT}. \quad (10.5)$$

Из соотношений (10.2)–(10.5) видно, что $\Psi < 0$ при изгибе зон вверх и $\Psi > 0$ при изгибе зон вниз. Если объемный Ψ_B и поверхностный Ψ_s электростатические потенциалы имеют одинаковые знаки, то приповерхностный слой обогащен основными носителями заряда, если же Ψ_B и Ψ_s имеют разные знаки, то имеют место слои обеднения или инверсии; последний будет иметь место в том случае, когда электростатический потенциал Ψ_s меняет знак и концентрация неосновных носителей заряда на поверхности больше, нежели основных (см рис. 10.1 и 10.4).

Уравнение Пуассона имеет вид:

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho(x)}{\epsilon_r \epsilon_0}, \quad (10.6)$$

где ϵ_r , ϵ_0 — относительная диэлектрическая проницаемость полупроводника и электрическая постоянная соответственно.

Полагая, что при всех значениях $\varphi(x)$ примеси N_a и N_d являются полностью ионизированными и равномерно распределены по всему объему полупроводника, для $\rho(x)$ при произвольном значении x можно записать:

$$\rho(x) = -e(n - p + N_a - N_d). \quad (10.7)$$

Принимая во внимание условие электронейтральности для объема полупроводника

$$n_0 - p_0 + N_a - N_d = 0, \quad (10.8)$$

выражение (10.7) можно трансформировать в

$$\rho(x) = -e [n(n - n_0) - (p - p_0)]. \quad (10.9)$$

Используя (10.2) и (10.3), можно получить $\rho(x)$ в функции поверхностного электростатического потенциала:

$$\rho(x) = -e [n_0 (e^{e\varphi/kT} - 1) - p_0 (e^{-e\varphi/kT} - 1)]. \quad (10.10)$$

Введем обозначения

$$Y = \frac{e\varphi}{kT}; \quad \lambda = \frac{n_0}{n_i} = \frac{n_i}{p_0} = e^{\Psi_B/kT}; \quad L_d^2 = \frac{\epsilon_r \epsilon_0 kT}{2e^2 n_i}. \quad (10.11)$$

Величина Y есть безразмерный электростатический потенциал; Y отрицателен при изгибах зон вверх и положителен при изгибах зон вниз; λ есть характеристика степени легирования полупроводника; $\lambda > 1$ для полупроводника n -типа, этому соответствует $\Psi_B > 0$; $\lambda < 1$ для полупроводника p -типа, объемный электростатический потенциал является величиной отрицательной; L_d по физическому смыслу есть дебаевский радиус экранирования для собственного полупроводника.

Используя обозначения (10.11) и подставляя (10.10) в (10.6), получаем:

$$2 \frac{d^2 Y}{dx^2} = L_d^{-2} [\lambda (e^Y - 1) - \lambda^{-1} (e^{-Y} - 1)]. \quad (10.12)$$

Граничные условия для решения (10.12) следующие: при $x = 0$

$$Y = Y_s, \quad (10.13)$$

а при $x \rightarrow \infty$

$$Y = 0 \text{ и } \frac{dY}{dx} = 0. \quad (10.14)$$

Первый интеграл (10.12) находим, используя тождество

$$\frac{d}{dx} \left(\frac{dY}{dx} \right)^2 = 2 \frac{d^2 Y}{dx^2} \frac{dY}{dx}. \quad (10.15)$$

Умножая правую и левую части уравнения (10.12) на $\frac{dY}{dx}$ и интегрируя, а затем извлекая квадратный корень, получаем:

$$\frac{dY}{dx} = \pm L_d^{-1} F(\lambda, Y) + C, \quad (10.16)$$

где

$$F(\lambda, Y) = [\lambda (e^Y - 1) - \lambda^{-1} (e^{-Y} - 1) + (\lambda^{-1} - \lambda)]^{1/2}. \quad (10.17)$$

Для отрицательных значений Y $\frac{dY}{dx} > 0$ и, следовательно, перед $F(\lambda, Y)$ необходимо выбрать плюс, а для положительных значений Y перед $F(\lambda, Y)$ необходимо выбрать знак минус. Использование граничного условия (10.14) приводит к $C = 0$. Для энергетической диаграммы рис. 10.4 $Y < 0$, поэтому

$$\frac{dY}{dx} = L_d^{-1} F(\lambda, Y). \quad (10.18)$$

Определим полный положительный заряд Q_{sp} в приповерхностной области полупроводника. Используя соотношение для L_d и (10.18), будем иметь:

$$Q_{sp} = \int_0^{\infty} \rho(x) dx = -\frac{\epsilon_r \epsilon_0 kT}{e} \int_0^{\infty} \frac{d^2 Y}{dx^2} dx = \frac{\epsilon_r \epsilon_0 kT}{e} \frac{dY}{dx} \Big|_{x=0} = 2en_i L_d F(\lambda, Y_s). \quad (10.19)$$

Таким образом, заряд в приповерхностной области Q_{sp} определяется значением поверхностного потенциала Y_s и уровнем легирования полупроводника (рис. 10.5).

В ходе зависимости $Q_{sp}(Y_s)$, определяемой из (10.17), можно выделить следующие характерные участки. Для определенности будем полагать, что полупроводник n -типа, т. е. $\lambda \gg 1$, и достаточно сильно легированный, так что $\lambda \gg \lambda^{-1}$.

1. При больших положительных значениях Y в соотношении (10.17) преобладающим является член λe^Y , знак Q_{sp} отрицательный и, следовательно, концентрация основных носителей заряда в приповерхностной области полупроводника выше, чем в объеме ($n_s > n_0$).

Это режим обогащения (см. рис. 10.2).

2. При $Y = 0$ $Q_{sp} = 0$, а значит, изгиб зон отсутствует. Это случай «плоских зон».

3. При малых отрицательных Y [диапазон изменений Y определяется условием $(1+Y)\lambda > > \lambda^{-1}e^{-Y}$] скорость убывания концентрации основных носителей заряда — электронов больше скорости нарастания концентрации дырок, вследствие чего приповерхностный слой полупроводника обедняется подвижными носителями заряда. Это режим обеднения (область I, рис. 10.1, а). Положительный пространственный заряд в этом случае создается в основном ионами донорной примеси.

Если вспомнить, что при выводе всех соотношений полагалась полная

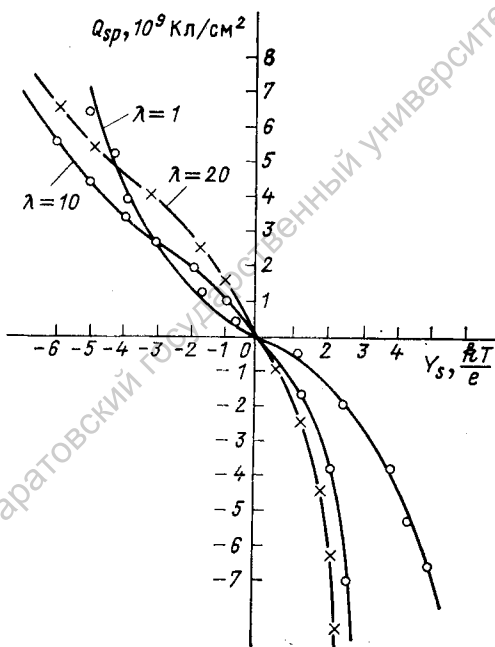


Рис. 10.5. Зависимость заряда в приповерхностном слое от безразмерного электростатического потенциала при различных уровнях легирования полупроводника

ионизация примеси при любых значениях Y , то увеличение Q_{sp} в этом диапазоне означает одновременное увеличение ширины области пространственного заряда, которое происходит до тех пор, пока преимущественный вклад в Q_{sp} не будут вносить дырки. Под действием поля, соответствующего отрицательным значениям Y (направление этого поля из полупроводника к поверхности), подвижные дырки формируют слой положительного объемного заряда, непосредственно прилегающий к поверхности. Образуется инверсионный слой (область II, рис. 10.1, а).

4. Условием образования инверсионного слоя является преобладание члена $\lambda^{-1}e^{-Y}$ над всеми остальными членами в (10.17). Здесь заряд Q_{sp} , как и в случае обогащенного слоя, экспоненциально возрастает с ростом Y .

График зависимости пространственного заряда Q_{sp} от безразмерного электростатического потенциала Y_s , рассчитанной с использованием (10.17), представлен на рис. 10.5 при λ в качестве параметра.

Как уже отмечалось, в случае свободной поверхности и отсутствия внешних электрических полей заряд в области пространственного заряда Q_{sp} равен и противоположен по знаку заряду в поверхностных состояниях Q_{ss} . Величина и знак Q_{ss} определяются характером поверхностных состояний (акцепторные или донорные), числом состояний на каждом уровне и положением уровня Ферми на поверхности. В принятых обозначениях на рис. 10.4 $e\Psi_s = F - F_i$. Если считать, что E_s есть энергия поверхностного уровня, отсчитанная от F_i , то поверхностный заряд Q_{ss} для полупроводника, имеющего, например, поверхностные акцепторные уровни E_s с концентрацией N_s , равен:

$$Q_{ss} = - \frac{eN_s}{e^{(E_s - e\Psi_s)/kT} + 1}. \quad (10.20)$$

Следовательно, используя (10.18) и (10.20), можно записать:

$$Q_{sp} = 2en_i L_d F(\lambda, Y_s) = -Q_{ss} = \frac{eN_s}{e^{(E_s - e\Psi_s)/kT} + 1}. \quad (10.21)$$

Выражение (10.21) можно использовать для нахождения электростатического потенциала свободной поверхности, если известны уровень легирования полупроводника λ , концентрация N_s и энергия E_s поверхностного уровня. При излучении поверхностных свойств полупроводников чаще приходится решать обратную задачу — по известному значению Q_{ss} в функции Y_s ищутся (а точнее, подбираются) N_s и F_s . Для нахождения $Q_{ss}(Y_s)$ используется наиболее распространенный метод исследования параметров поверхности полупроводников — метод эффекта поля.

Для получения вида зависимости Y от x необходимо проинтегрировать (10.18) с граничными условиями $Y = Y_s$ при $x = 0$ и $Y = 0$ при $x \rightarrow \infty$:

$$\int_{Y_s}^0 \frac{dY}{\mp F(\lambda, Y)} = \frac{x}{L_d}. \quad (10.22)$$

Интеграл в (10.22) при $F(\lambda, Y)$, в общем виде определяемый соотношением (10.17), не может быть вычислен в квадратурах. Поэтому для выяснения хода зависимости рассматривают частные случаи, в которых возможно получить явный вид зависимости Y от x .

10.3. ЭФФЕКТ ПОЛЯ

Из предыдущего параграфа следует, что изменением поверхностного электростатического потенциала можно в широких пределах изменять концентрацию подвижных носителей заряда в слое пространственного заряда полупроводника. Обозначим через ΔN и ΔP избыток поверхностных концентраций электронов и дырок, которые определим как

$$\Delta N = \int_0^{\infty} [n(x) - n_0] dx; \quad (10.23)$$

$$\Delta P = \int_0^{\infty} [p(x) - p_0] dx. \quad (10.24)$$

Значения ΔN и ΔP могут быть вычислены в функции поверхностного электростатического потенциала Y_s с использованием (10.18) следующим образом:

$$\begin{aligned} \Delta N &= \int_0^{\infty} [n(x) - n_0] dx = n_0 \int_0^{\infty} (e^Y - 1) dx = \\ &= n_0 \int_{Y_s}^0 (e^Y - 1) \left(\frac{dY}{dx} \right)^{-1} dY = n_0 L_d \int_0^{Y_s} \frac{e^Y - 1}{F(\lambda, Y)} dY; \end{aligned} \quad (10.25)$$

$$\begin{aligned} \Delta P &= \int_0^{\infty} [p(x) - p_0] dx = p_0 \int_0^{\infty} (e^{-Y} - 1) dx = \\ &= p_0 \int_{Y_s}^0 (e^{-Y} - 1) \left(\frac{dY}{dx} \right)^{-1} dY = -p_0 L_d \int_{Y_s}^0 \frac{e^{-Y} - 1}{F(\lambda, Y)} dY = \\ &= p_0 L_d \int_0^{Y_s} \frac{e^{-Y} - 1}{F(\lambda, Y)} dY. \end{aligned} \quad (10.26)$$

Избыточные электроны и дырки могут двигаться вдоль поверхности при приложении электрического поля, обуславливая тем самым изменение проводимости образца. Вызванная наличием ΔN и ΔP избыточная проводимость носит название *поверхностной проводимости*. Если предположить, что значение подвижностей электронов и дырок в области пространственного заряда

равно их значению в объеме полупроводника, то для поверхностной проводимости можно написать:

$$G_s = e\mu_n \Delta N + e\mu_p \Delta P. \quad (10.27)$$

Поскольку ΔN и ΔP являются функциями уровня легирования и поверхностного электростатического потенциала Y_s , то и G_s является функцией тех же величин. Однако характер этой зависимости не столь уже очевиден, если ΔN и ΔP определяется (10.25) и (10.26). Интегралы, входящие в (10.25) и (10.26), табулированы при различных Y и λ , что облегчает задачу вычисления G_s по (10.27).

На рис. 10.6 представлена зависимость $G_s(Y_s)$ с λ в качестве параметра, где $\lambda > 1$, что соответствует полупроводнику n -типа. На рисунке выделены характерные участки этой зависимости.

При $Y_s = 0$ — случай «плоских зон» и поверхностная проводимость $G_s = 0$.

Для положительных значений $Y_s > 0$ (изгиб зон «вниз», см. рис. 10.2) концентрация основных носителей заряда в области пространственного заряда растет, поверхностная проводимость $G_s > 0$ и монотонно возрастает с ростом Y_s (правая часть рис. 10.6). Это соответствует режиму обогащения.

При отрицательном Y_s поверхностная проводимость $G_s < 0$ и с ростом $|Y_s|$ вплоть до $Y_s = 2 e\psi_B/kT$ уменьшается, поскольку концентрация подвижных носителей заряда в слое пространственного заряда меньше, чем в случае $Y_s = 0$. Это соответствует режиму обеднения. Эта ситуация будет иметь место до тех пор, пока скорость нарастания неосновных носителей заряда — в данном случае дырок — с изменением Y_s не станет больше, нежели скорость убывания концентрации основных носителей заряда — электронов. Это соответствует условию

$$n_0 \approx p_0, \quad (10.28)$$

т. е. концентрация электронов в объеме полупроводника равна концентрации дырок на поверхности.

Значение поверхностной проводимости, соответствующее этому условию, будет минимально и может быть найдено из условия экстремума

$$\frac{dG_s}{dY_s} = e \left[\mu_n n_0 L_d \frac{e^{Y_s} - 1}{F(\lambda, Y_s)} + \mu_p p_0 L_d \frac{e^{-Y_s} - 1}{F(\lambda, Y_s)} \right] = 0. \quad (10.29)$$

При достаточно высоком уровне легирования $\left| \frac{e\psi_B}{kT} \right| \gg 1$ в нашем случае минимум G_s будет наблюдаться при достаточно больших по

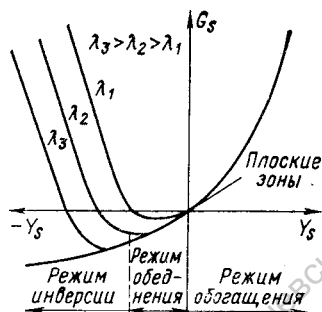


Рис. 10.6. Приповерхностная проводимость в функции безразмерного электростатического потенциала при различных уровнях легирования полупроводника

абсолютному значению поверхностных электростатических потенциалов Y_s и, следовательно, выражение (10.29) может быть упрощено при использовании неравенств $e^{Y_s} < 1$ и $e^{-Y_s} > 1$ для полупроводника n -типа и $e^{Y_s} > 1$ и $e^{-Y_s} < 1$ для полупроводника p -типа (следует помнить о знаке Y_s). С учетом этого (10.29) примет вид:

$$\mu_p p_0 e^{-Y_s} - \mu_n n_0 = 0. \quad (10.30)$$

Поскольку $n_0 = n_i e^{e\varphi_B/kT}$ и $p_0 = n_i e^{-e\varphi_B/kT}$,

$$\frac{\mu_n}{\mu_p} e^{2e\varphi_B/kT} = e^{-Y_s}. \quad (10.31)$$

При $\mu_n = \mu_p$ условие (10.31) соответствует (10.28), последнее обычно используют в качестве критерия образования инверсионного слоя. При этом выражение (10.28) тем точнее, чем лучше для данного полупроводника выполняется условие $\mu_n = \mu_p$.

Последующее увеличение $|Y_s|$ приводит к резкому увеличению концентрации дырок, образованию инверсионного слоя и, следовательно, к увеличению поверхностной проводимости G_s . Все эти участки кривой $G_s(Y_s)$ приведены на рис. 10.6.

До этого полагалось, что значения подвижностей носителей заряда в объеме и в слое пространственного заряда вблизи поверхности равны. Однако это не так в том случае, когда при соударении с поверхностью носитель заряда полностью или частично теряет дрейфовую составляющую скорости. Это так называемое диффузное рассеяние. Это так называемое диффузное рассеяние приводит к уменьшению подвижности носителя заряда, движущегося в потенциальной яме (основного носителя заряда в случае слоев обогащения и неосновного в случае слоев инверсии). Физической причиной снижения подвижности является то, что при появлении электрического поля, нормального к поверхности, увеличивается составляющая скорости к поверхности и, как результат этого, уменьшается время соударения носителя заряда с поверхностью. На рис. 10.7 представлена зависимость подвижности носителей заряда в слое пространственного заряда от поверхностного электростатического потенциала в случае полностью диффузного рассеяния.

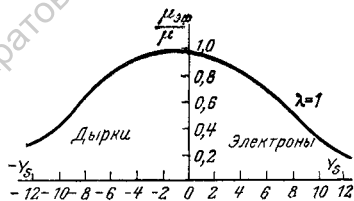


Рис. 10.7. Влияние поверхностного рассеяния на подвижность носителей заряда в области поверхностного заряда

Наиболее эффективным способом управления значением поверхностного электростатического потенциала (а следовательно, и значением поверхностной проводимости G_s и емкости пространственного заряда полупроводника G_{sp}) является приложение к полупроводнику электрического поля, нормального к поверхности полупроводника,

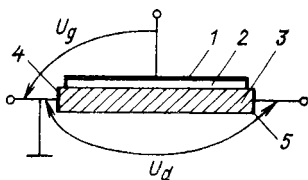


Рис. 10.8. Схематическое изображение структуры, используемой для измерения эффекта поля:

1 — металлический электрод (затвор); 2 — диэлектрик; 3 — полупроводник; 4 — омический контакт (исток); 5 — омический контакт (сток)

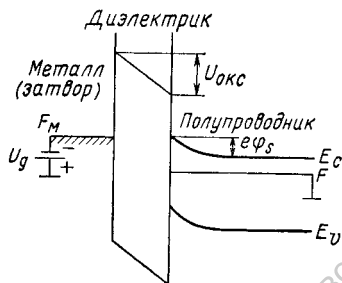


Рис. 10.9. Энергетическая диаграмма структуры, изображенной на рис. 10.8

Это сравнительно просто достигается в структуре, где полупроводник служит одной из обкладок конденсатора, отделенной от второй обкладки — металлического электрода — слоем диэлектрика (так называемая МДП-структура). *Эффектом поля* называют явление изменения проводимости полупроводникового образца под действием электрического поля, нормального к его поверхности. На рис. 10.8 представлена МДП-структура, используемая обычно при измерениях эффекта поля, на рис. 10.9 — соответствующая ей энергетическая диаграмма.

Экспериментально снимается зависимость изменения проводимости образца при вариации значения потенциала U_g , приложенного к металлическому электроду. Из диаграммы, представленной на рис. 10.9, следует:

$$U_g = U_{\text{окс}} + \varphi_s, \quad (10.32)$$

где $U_{\text{окс}}$ — падение потенциала в диэлектрике; $\varphi_s = \frac{kT}{e} Y_s$.

Если обозначить емкость диэлектрика через $C_{\text{окс}} = \frac{\epsilon_r \epsilon_0 n_0}{d}$, где d — толщина окисла, $\epsilon_{\text{окс}}$ — относительная диэлектрическая проницаемость окисла, то индуцированный в полупроводник заряд $Q_{\text{инд}} = C_{\text{окс}} U_{\text{окс}}$.

В простейшем случае, когда контактной разностью потенциалов между металлом и полупроводником можно пренебречь в сравнении с U_g , а на границе раздела полупроводник — диэлектрик и в объеме диэлектрика отсутствуют состояния, способные захватить носители заряда, весь индуцированный заряд будет участвовать в создании дополнительной проводимости. Тогда можно записать:

$$Q_{\text{инд}} = C_{\text{окс}} U_{\text{окс}} = C_{\text{окс}} (U_g - \varphi_s) = 2en_i L_d F(\lambda, Y_s). \quad (10.33)$$

Из (10.33) можно найти φ_s , соответствующий данному значению U_g , и, следовательно, значение G_s в функции Y_s . При выполнении вышеуказанных условий экспериментальная кривая $G_s(U_g)$ дает возможность при известном знаке U_g однозначно определить значение электростатического потенциала. Когда проводимость приповерх-

ностного слоя носит монополярный характер (это имеет место в ярко выраженных слоях обогащения и инверсии), можно ввести в рассмотрение подвижность при эффекте поля $\mu_{э, п}$, которую определяют как отношение приращения проводимости к полному индуцированному заряду:

$$\mu_{э, п} = \frac{dG_s}{dQ_{инд}} = \frac{d(e \Delta N \mu_n)}{d(e \Delta N)}. \quad (10.34)$$

Здесь принято во внимание, что при $|Y_s| \gg 1$ $Q_{инд} = C_{окс} U_{окс} \approx e \Delta N$. При отсутствии явлений захвата $\mu_{э, п}$ равна омической подвижности носителей заряда в слое пространственного заряда.

В опытах по эффекту поля обычно стремятся выполнить условие $C_{окс} \ll C_{сп}$, что дает возможность индуцированный заряд определять как $Q_{инд} = C_{окс} U_q$ (т. е. $U_q \gg \phi_s$) и экспериментальную зависимость строить в координатах $G_s(Q_{инд})$. Теоретическая зависимость в тех же координатах при известных подвижностях электронов и дырок и уровне легирования полупроводника легко строится по схеме

$$C_{окс} U_q \rightarrow Q_{инд}(U_s) \rightarrow U_s \rightarrow G_s(Y_s).$$

Реально наблюдаемая в экспериментах по эффекту поля зависимость поверхностной проводимости от индуцированного заряда существенно отличается от теоретической. Причины этого состоят в следующем: 1) между полупроводником и металлом существует контактная разность потенциалов $e\phi_k = \Phi_m - \Phi_n$; 2) на границе раздела полупроводник — диэлектрик и в самом диэлектрике практически всегда существуют состояния, способные нести нескомпенсированный заряд определенного знака и не меняющие своего зарядового состояния при всех достижимых в данном эксперименте значениях поверхностного электростатического потенциала Y_s ; 3) имеются такие состояния, заполнение которых электронами зависит от значения Y_s . Первые две из указанных причин приводят к тому, что в полупроводнике существует поверхностная проводимость ($G_s \neq 0$) в отсутствие напряжения на металлическом электроде ($U_q = 0$) и, следовательно, происходит параллельный сдвиг экспериментальной кривой $G_s(Q_{инд})$ относительно теоретической вдоль оси зарядов на значение Q_{s0} . Влияние третьей из указанных причин приводит к гораздо более существенной деформации экспериментальной кривой: кроме сдвига наблюдается изменение формы кривой (рис. 10.10).

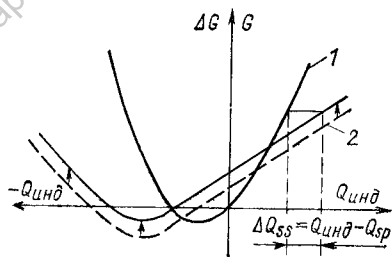


Рис. 10.10. Поверхностная проводимость в зависимости от индуцированного в полупроводнике заряда (стрелки поясняют методику расчета параметров поверхности из эксперимента):

1 — теоретическая кривая; 2 — экспериментальная кривая

Таким образом, количественное и качественное различия между теоретической и экспериментальной кривыми поверхностной проводимости от индуцированного заряда позволяет судить о концентрации и энергетическом спектре поверхностных состояний в структуре полупроводник — диэлектрик. Это обычно и является основной задачей метода эффекта поля. При наличии поверхностных состояний лишь часть индуцированного в полупроводник заряда Q_{sp} участвует в создании поверхностной проводимости, другая часть, т. е. Q_{ss} , захватывается на поверхностные состояния:

$$Q_{\text{инд}} = C_{\text{окс}} U_{\text{окс}} = \mp Q_{sp} \mp Q_{ss} = \\ = \mp 2en_i L_d F(\lambda, Y_s) \mp e \sum_i N_{si} f_{si}(E_{si}), \quad (10.35)$$

где N_{si} — число состояний на уровне захвата с энергией E_{si} ; f_{si} — соответствующая функция распределения (знак минус во втором слагаемом относится к акцепторным, плюс — к донорным состояниям).

В дальнейшем вопросы кинетики захвата носителей заряда на поверхностные состояния будут рассмотрены в § 10.4, здесь лишь отметим, что предполагаются такие условия измерения экспериментальной кривой $G_s(Q_{\text{инд}})$, при которых значения f_{si} однозначно определяются значением поверхностного потенциала ϕ_s (или уровня Ферми на поверхности), т. е. все уровни поверхностных состояний находятся в равновесии друг с другом и зоной проводимости и валентной зоной (так называемый квазиравновесный эффект поля).

Эффективность управления значениями Y_s и G_s с помощью поперечного электрического поля зависит от соотношения между изменением заряда в приповерхностной области ΔQ_{sp} и захваченного заряда ΔQ_{ss} при заданном ΔU_g ; чем больше ΔQ_{ss} , тем меньше ΔY_s и ΔG_s . Поскольку при заданном знаке $dG_s/dQ_{\text{инд}} = \mu_{s,п}$ G_s определяется только значением Y_s , то очевидно, что при различном захваченном заряде Q_{ss} необходимо индуцировать в полупроводник различный заряд для достижения одного и того же значения G_s (или Y_s). Разность между индуцированными зарядами для достижения данного значения G_s (или Y_s) в экспериментальной и теоретической кривой дает значение захваченного заряда для заданного Y_s . Эту разность, как правило, и определяют при исследовании по методу эффекта поля.

Вышесказанное предполагает, что для расчета необходимо построить теоретическую кривую $G_s(Q_{\text{инд}})$, для чего должны быть известны параметры полупроводника — собственная концентрация n_i , уровень легирования $N_d(N_a)$, подвижности электронов μ_n и дырок μ_p с учетом характера поверхностного рассеяния, а также емкость диэлектрика $C_{\text{окс}}$. Построение может быть проведено с помощью соотношений (10.25), (10.26) и (10.27). Обработка данных по эффекту поля возможна в том случае, когда экспериментальная кривая $G_s(Q_{\text{инд}})$ имеет подобно теоретической минимум. Это происходит тогда, когда плотность поверхностных состояний, способных менять заряд в условиях измерений, относительно мала, а напряженность поперечного

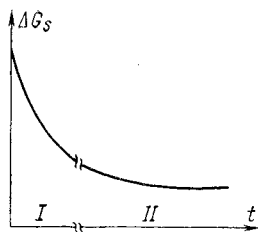


Рис. 10.11. Кривая релаксации поверхностной проводимости: *I* — участок быстрой релаксации (10^{-3} – 10^{-8} с); *II* — участок медленной релаксации (более 10^{-2} с)

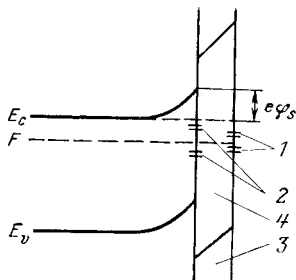


Рис. 10.12. Быстрые и медленные состояния на поверхности германия:

1 — медленные состояния; *2* — быстрые состояния; *3* — слой окиси германия; *4* — запрещенная зона окиси

поля достаточно велика. Значение G_s , соответствующее минимуму как экспериментальной, так и теоретической кривых $G_s(Q_{\text{инд}})$, определяется условием (10.31) и может быть принято за точку отсчета абсолютных значений Y_s . Значение поверхностной проводимости также удобно отсчитывать от ее значения в минимуме. Для этого минимум экспериментальной кривой совмещают с минимумом теоретической: этапы расчета ясны из рис. 10.10.

Данные исследования эффекта поля на недостаточно качественно приготовленных образцах МДП-структур на постоянном токе показали, что в процессе его измерения после приложения постоянного напряжения к управляющему электроду поверхностная проводимость релаксирует к нулю при всех возможных значениях U_g любого знака. Это свидетельствует о большом количестве акцепторных и донорных состояний на границе раздела полупроводник — диэлектрик, захватывающих индуцированный заряд и экранирующих объем полупроводника от проникновения электрического поля. Более детальное изучение кинетики эффекта поля показало, что существуют два участка спада на кривой поверхностной проводимости, характеризующихся резко различными постоянными спада (рис. 10.11). Это позволило говорить о двух типах поверхностных состояний с различными временами захвата. К первому типу состояний относятся так называемые «быстрые состояния», которые характеризуются временами захвата около 10^{-8} – 10^{-3} с; полагают, что они расположены непосредственно на границе раздела полупроводник — диэлектрик. Ко второму типу состояний относятся «медленные состояния», расположенные на внешней стороне окисла и потому характеризующиеся большими временами захвата — от 10^{-3} с до нескольких часов (рис. 10.12). Применение в измерениях синусоидального напряжения определенной частоты на управляющем электроде позволяет исключить влияние «медленных» поверхностных состояний на результаты измерений в эффекте поля.

10.4. СКОРОСТЬ ПОВЕРХНОСТНОЙ РЕКОМБИНАЦИИ

Из большого числа состояний, возникающих на поверхности полупроводника, некоторые состояния могут 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 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. Для характеристики процесса рекомбинации на поверхности вводят величину s , называемую скоростью поверхности 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 — средняя 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

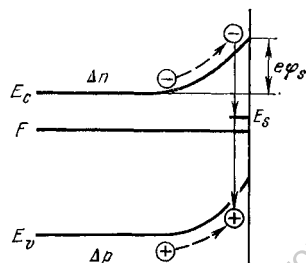


Рис. 10.13. Рекомбинация через поверхностные уровни

$$\frac{1}{4} v_0 p_0 = \frac{1}{4} r v_0 p_0 + q_{s0}, \quad (10.36)$$

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 разницей 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):

$$-\left(\frac{dp}{dt}\right)_r = R_p = \frac{\gamma_{ns}\gamma_{ps}N_s(n_s p_s - n_1 p_1)}{\gamma_{ns}(n_s + n_1) + \gamma_{ps}(p_s + p_1)}, \quad (10.41)$$

где γ_{ns} и γ_{ps} — вероятность захвата поверхностным уровнем электронов и дырок из соответствующих зон в течение одной секунды, см³/с; n_s и p_s — неравновесные концентрации электронов и дырок на поверхности полупроводника; n_1 и p_1 — 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 $-\left(\frac{\partial p}{\partial t}\right)_r = -\left(\frac{\partial n}{\partial t}\right)_r$ или $R_p = R_n = R$, то имеют место равные 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 рекомбина-

ции на поверхности и в объеме полупроводника. 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 уравниваются друг друга. В этом случае концентрации носителей заряда будут определяться распределением Больцмана. Если изгиб зоны у поверхности $e\varphi_s$, а n и p — концентрации неравновесных электронов и дырок в объеме полупроводника, то их поверхностные значения

$$n_s = n e^{e\varphi_s/kT} = n e^{Y_s}; \quad p_s = p e^{-e\varphi_s/kT} = p e^{-Y_s}, \quad (10.42)$$

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

$$R = \frac{N_s \gamma_{ns} \gamma_{ps} (np - n_0 p_0)}{\gamma_{ns} (n_s + n_1) + \gamma_{ps} (p_s + p_1)}. \quad (10.44)$$

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 $\Delta n, \Delta p \ll n_0, p_0$, получаем

$$R = \frac{N_s \gamma_{ns} \gamma_{ps} (n_0 + p_0) \Delta p}{\gamma_{ns} (n_{s0} + n_1) + \gamma_{ps} (p_{s0} + p_1)} = \Delta q_s = s \Delta p. \quad (10.45)$$

Используя (10.43), получаем соотношение для скорости поверхностной рекомбинации в зависимости от поверхностного электростатического потенциала Y_s

$$s = \frac{N_s \gamma_{ns} \gamma_{ps} (n_0 + p_0)}{\gamma_{ns} (n_0 e^{Y_s} + n_1) + \gamma_{ps} (p_0 e^{-Y_s} + p_1)}. \quad (10.46)$$

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

$$Y_{max} = \frac{1}{2} \ln \frac{\gamma_{ns} n_0}{\gamma_{ps} p_0}. \quad (10.47)$$

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.

Вид зависимости $s(Y_s)$ в соответствии с (10.46) представлен на рис. 10.14. Из (10.46) очевидно, что если из эксперимента известно

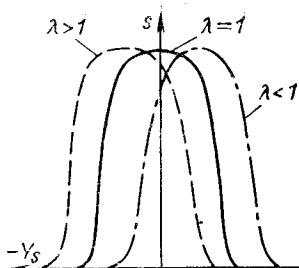


Рис. 10.14. Зависимость скорости поверхностной рекомбинации от безразмерного поверхностного электростатического потенциала

значение электростатического потенциала, соответствующее максимуму скорости поверхностной рекомбинации, то можно вычислить отношение вероятностей захвата электрона и дырки рекомбинационным уровнем или, поскольку $\gamma_{ns, ps} = \sigma_{n, p} u_0 N_s$, 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.

10.5. ВЛИЯНИЕ ПОВЕРХНОСТНОЙ РЕКОМБИНАЦИИ НА ВРЕМЯ ЖИЗНИ НОСИТЕЛЕЙ ЗАРЯДА В ОБРАЗЦАХ КОНЕЧНЫХ РАЗМЕРОВ

Наличие поверхностных рекомбинационных ловушек, естественно, приводит к увеличению общего числа рекомбинационных центров в образце и, следовательно, к уменьшению времени жизни носителей заряда.

Рассмотрим распределение избыточной концентрации носителей заряда вблизи поверхности, на которой $s \neq 0$ в случае, когда генерация избыточных электронно-дырочных пар производится равномерно по всему объему образца со скоростью G . Для определенности выбираем полупроводник n -типа, координата x нормальна к поверхности образца и возрастает от поверхности в глубь образца, внешнее электрическое поле отсутствует.

Избыточная концентрация неосновных носителей — дырок $\Delta p(x)$ может быть получена из решения стационарного уравнения непрерывности

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial J_p}{\partial x} - \frac{\Delta p}{\tau_p} + G = 0. \quad (10.48)$$

Поскольку напряженность электрического поля в области пространственного заряда невелика, дрейфовую составляющую тока можно не учитывать. Тогда $J_p = -eD_p \frac{dp}{dx}$ и уравнение (10.48) примет вид:

$$D_p \frac{d^2 \Delta p}{dx^2} - \frac{\Delta p}{\tau_p} + G = 0. \quad (10.49)$$

Граничное условие в точке $x = 0$ для потока дырок к поверхности может быть задано в виде

$$D_p \frac{d\Delta p}{dx} = -\frac{J_p}{e} = \Delta q_s = s \Delta p. \quad (10.50)$$

Поскольку потоки электронов и дырок, обусловленные поверхностной рекомбинацией, носят чисто диффузионный характер, то можно ожидать, что влияние поверхности на распределение избыточной концентрации дырок будет сказываться на расстояниях по-

рядка нескольких диффузионных длин от поверхности. Следовательно, для $x \rightarrow \infty$ уравнение непрерывности (10.49) будет:

$$-\frac{\Delta p}{\tau_p} + G = 0 \text{ и } \Delta p|_{x \gg L_p} = G\tau_p. \quad (10.51)$$

Общее решение (10.49) имеет вид:

$$\Delta p(x) = Ae^{-x/L_p} + G\tau_p. \quad (10.52)$$

Для нахождения A используем граничное условие (10.50) в точке $x = 0$

$$-\frac{A}{L_p}D_p = s(A + G\tau_p). \quad (10.53)$$

Отсюда

$$A = -\frac{Gs\tau_p L_p}{D_p + sL_p} = -\frac{Gs\tau_p^2}{L_p + s\tau_p}. \quad (10.54)$$

Окончательно имеем:

$$\Delta p(x) = G\tau_p \left[1 - \frac{s\tau_p}{L_p + s\tau_p} e^{-x/L_p} \right] = G\tau_p \left[1 - \frac{s/v_D}{1 + s/v_D} e^{-x/L_p} \right], \quad (10.55)$$

где $v_D = L_p/\tau_p$ — диффузионная скорость.

Физический смысл полученного соотношения достаточно очевиден:

1) если $s \ll v_D$, то для обеспечения малого диффузионного потока на поверхностную рекомбинацию необходимы малые градиенты концентраций избыточных носителей заряда, т. е. Δp мало меняется с изменением x и $\Delta p \approx G\tau_p$;

2) при $s \gg v_D$ будет иметь место сильная зависимость Δp от x , поскольку необходимо иметь достаточный диффузионный поток на поверхностную рекомбинацию. Характер зависимости $\Delta n = \Delta p$ от x при различных соотношениях s и v_D иллюстрируется рис. 10.15.

Концентрацию избыточных дырок на поверхности полупроводника Δp_s можно определить, если в (10.55) положить $x = 0$. Она будет равна:

$$\Delta p_s = \frac{G\tau_p L_p}{L_p + s\tau_p}. \quad (10.56)$$

Введем $\tau_s = \Delta p_s/G$, которое является временем жизни неосновных носителей заряда на поверхности полупроводника. Из (10.56) следует, что

$$\tau_s = \tau_p \frac{L_p}{L_p + s\tau_p}. \quad (10.57)$$

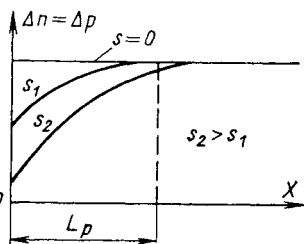


Рис. 10.15. Влияние скорости поверхностной рекомбинации на распределение избыточной концентрации носителей заряда

Тогда

$$1 - \frac{\tau_s}{\tau_p} = \frac{s\tau_p}{L_p + s\tau_p}. \quad (10.58)$$

В таком случае равенство (10.55) будет иметь вид:

$$\Delta p(x) = G\tau_p \left[1 - \left(1 - \frac{\tau_s}{\tau_p} \right) e^{-x/L_p} \right] = G\tau_p \left[1 + \left(\frac{\tau_s}{\tau_p} - 1 \right) e^{-x/L_p} \right]. \quad (10.59)$$

Отсюда получим время жизни неосновных носителей заряда в приповерхностной области полупроводника:

$$\tau(x) = \frac{\Delta p(x)}{G} = \tau_p \left[1 + \left(\frac{\tau_s}{\tau_p} - 1 \right) e^{-x/L_p} \right]. \quad (10.60)$$

ГЛАВА ОДИННАДЦАТАЯ

ПОГЛОЩЕНИЕ СВЕТА ПОЛУПРОВОДНИКАМИ

11.1. СПЕКТР ОТРАЖЕНИЯ И СПЕКТР ПОГЛОЩЕНИЯ

Допустим, что на пластину полупроводника падает пучок монохроматического излучения интенсивностью I_0 (рис. 11.1). Вследствие процесса отражения и поглощения интенсивность света, прошедшего через образец, будет уменьшена. Если I_R — интенсивность отраженного света, то доля отраженного излучения составит величину, равную

$$R = I_R/I_0, \quad (11.1)$$

получившую название коэффициента отражения.

Зависимость коэффициента отражения вещества от энергии кванта падающего света $R(h\nu)$ или длины волны $R(\lambda)$ называют спектром отражения.

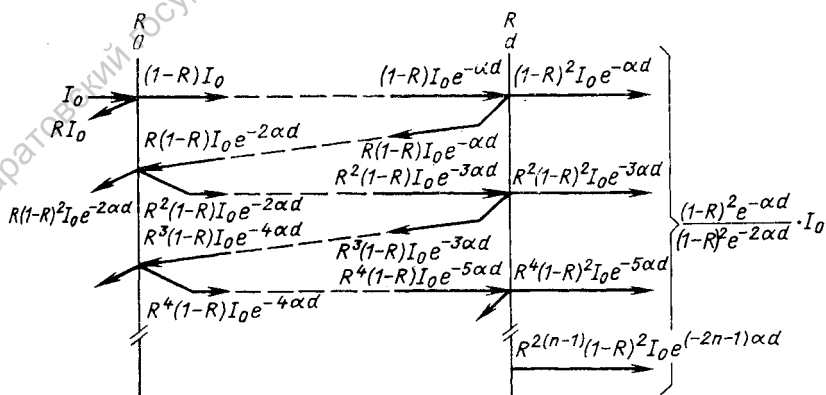


Рис. 11.1. Потoki энергии в системе с многократным внутренним отражением

С учетом отражения через первую поверхность образца пройдет излучение $(1 - R) I_0$. Вследствие поглощения света в слое dx интенсивность излучения I в единицу времени уменьшится на dI . Количество поглощенной энергии dI пропорционально количеству падающего на слой энергии I и толщине поглощающего слоя dx :

$$-dI = \alpha I dx. \quad (11.2)$$

Коэффициент пропорциональности α , выражающий собой количество поглощенной энергии веществом из пучка единичной интенсивности излучения в единицу времени в слое единичной длины, носит название коэффициента поглощения.

Проинтегрировав уравнение (11.2), найдем интенсивность излучения $(1 - R) I_0 e^{-\alpha d}$, достигшего второй поверхности пластины. Интенсивность света, вышедшего из образца, будет равна $(1 - R) \times (1 - R) I_0 e^{-\alpha d}$. Свет, отраженный внутрь образца, как следует из рис. 11.1, выйдет из него ослабленным. С учетом этих эффектов соотношение для коэффициента пропускания T , выражающего собой отношение интенсивности света $I = \frac{(1 - R)^2 e^{-\alpha d}}{1 - R^2 e^{-2\alpha d}} I_0$, прошедшего через образец толщиной d , к интенсивности падающего света I_0 , будет

$$T = \frac{(1 - R)^2 e^{-\alpha d}}{1 - R^2 e^{-2\alpha d}}. \quad (11.3)$$

Если αd велико, то можно пренебречь вторым членом в знаменателе (11.3). В этом случае интенсивность света, прошедшего через образец толщиной d , с учетом отражения будет равна:

$$I = (1 - R)^2 I_0 e^{-\alpha d}. \quad (11.4)$$

Здесь α является характеристикой поглощающей среды и зависит от длины волны излучения. Зависимость коэффициента поглощения от энергии падающего на вещество кванта света $\alpha(h\nu)$ или длины волны $\alpha(\lambda)$ называют спектром поглощения.

Пусть имеется N поглощающих центров. Обозначим через σ вероятность поглощения одного фотона одним поглощающим центром в единицу времени. Другими словами, σ есть эффективное сечение поглощения одного фотона в единицу времени. Эффективное сечение σ зависит от энергии фотона и природы поглощающих центров.

Если через l_ϕ обозначить среднюю длину свободного пробега фотона в поглощающей среде, то согласно (5.7) она равна:

$$l_\phi = (\sigma N)^{-1}, \quad (11.5)$$

а коэффициент поглощения

$$\alpha = 1/l_\phi = \sigma N \quad (11.6)$$

есть вероятность поглощения фотона на единичной длине.

Допустим, что в полупроводнике имеются поглощающие центры различной природы. Если центры поглощения N_i характеризуются эффективным сечением σ_i , то

$$\alpha_i(h\nu) = 1/l_{i\phi} = \sigma_i N_i. \quad (11.7)$$

Полный коэффициент поглощения вещества α будет:

$$\alpha(h\nu) = \sum_i \sigma_i N_i = \sum_i \alpha_i, \quad (11.8)$$

т. е. он равен сумме коэффициентов поглощения различными центрами. Следовательно, полный спектр поглощения вещества складывается из спектров поглощения различных поглощающих центров.

При взаимодействии электронов полупроводника с электромагнитным излучением должны выполняться два закона: закон сохранения энергии и закон сохранения квазиимпульса. Если электрон до взаимодействия с квантом света, имеющим энергию $h\nu = \hbar\omega$ и импульс $\hbar\eta$, обладал энергией E и квазиимпульсом $\mathbf{p}$, а после взаимодействия имел соответственно E' и $\mathbf{p}'$, то эти законы запишутся в виде

$$E' = E + \hbar\omega; \quad (11.9)$$

$$\mathbf{p}' = \mathbf{p} + \hbar\eta. \quad (11.10)$$

Поглощение излучения в полупроводниках может быть связано с изменением энергетического состояния свободных или связанных с атомами электронов, а также с изменением колебательной энергии атомов решетки. В связи с этим в полупроводниках различают пять основных типов оптического поглощения: собственное, экситонное, поглощение свободными носителями заряда, примесное и решеточноепоглощение.

11.2. СОБСТВЕННОЕ ПОГЛОЩЕНИЕ ПРИ ПРЯМЫХ ПЕРЕХОДАХ

Если при поглощении полупроводником кванта излучения имеет место возбуждение электронов из валентной зоны в зону проводимости, то такое поглощение называют собственным или фундаментальным. При изучении собственного поглощения полупроводника следует учитывать строение его энергетических зон. Известные в настоящее время полупроводники в соответствии с конфигурацией энергетических зон можно разделить на два основных вида. У первого из них минимум энергии в зоне проводимости, характеризуемый волновым вектором $\mathbf{k}_{min}$, и максимум энергии в валентной зоне, определяемый волновым вектором $\mathbf{k}_{max}$, расположены в одной и той же точке зоны Бриллюэна (обычно в точке $\mathbf{k} = 0$). Другими словами, у этих полупроводников $\mathbf{k}_{min} = \mathbf{k}_{max}$ (рис. 11.2, а). В качестве примера полупроводника с таким построением зон можно указать антимонид индия (см. рис. 2.25),

У второго вида веществ экстремумы зоны проводимости и валентной зоны находятся при различных $\mathbf{k}$, так что в этом случае $\mathbf{k}_{min} \neq \mathbf{k}_{max}$ (рис. 11.2, б). К последнему типу веществ относится большинство полупроводников, в том числе германий и кремний (см. рис. 2.21).

Переходы электронов через запрещенную зону будут происходить прежде всего между энергетическими состояниями, соответствующими максимуму валентной зоны и минимуму зоны проводимости,

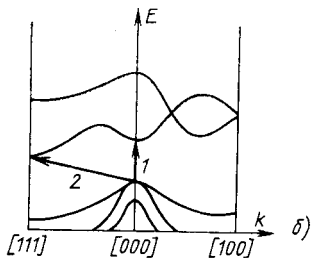
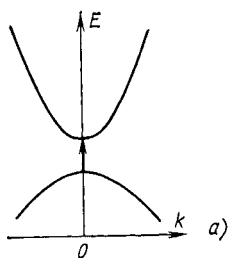


Рис. 11.2. Собственное поглощение при прямых межзонных переходах (а) и оптические переходы в полупроводнике со сложными энергетическими зонами (б)

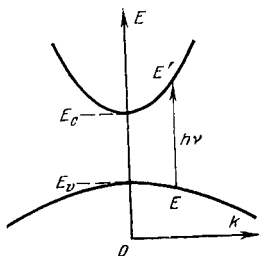


Рис. 11.3. Прямые переходы

т. е. при значениях квазиимпульса $\mathbf{p}$ или волнового вектора $\mathbf{k}$, близких к нулю, как это показано на рис. 11.2, а.

Импульс фотона $\hbar\mathbf{k} = 1/\lambda_\phi$ очень мал по сравнению с квазиимпульсом электрона, так как длина волны света λ_ϕ составляет 10^{-1} — 10^{-5} см, а длина волны электрона в кристалле при энергии электрона, соответствующей 300 К, равна примерно $5 \cdot 10^{-7}$ см. В связи с этим в (11.10) импульсом фотона можно пренебречь, поэтому

$$\mathbf{p}' = \mathbf{p} \quad (11.11)$$

или

$$\mathbf{k}' = \mathbf{k}. \quad (11.12)$$

Соотношение (11.11) или (11.12), называемое правилом отбора для электронных переходов, показывает, что в процессе взаимодействия электрона кристалла с полем излучения возможны только такие переходы, при которых волновой вектор электрона сохраняется. Эти переходы получили название в е р т и к а л ь н ы х, или п р я м ы х, п е р е х о д о в (рис. 11.2, а и переход 1 на рис. 11.2, б).

Рассмотрим собственное поглощение при прямых межзонных переходах для полупроводника, у которого зона проводимости и валентная зона обладают сферической симметрией. Будем полагать, что валентная зона целиком занята электронами, а зона проводимости полностью свободна. Это условие справедливо для нелегированных полупроводников при 0 К. Согласно рис. 11.3 каждому начальному состоянию E в валентной зоне при поглощении кванта света с энергией $h\nu$ отвечает такое конечное состояние E' в зоне проводимости, что

$$E' = h\nu + E. \quad (11.13)$$

В случае зон сферической симметрии в соответствии с (4.4) и (4.10)

$$E' = E_c + \frac{\hbar^2 k^2}{2m_n^*}; \quad E = E_v - \frac{\hbar^2 k^2}{2m_p^*}, \quad (11.14)$$

поэтому

$$\hbar v - E_g = \frac{\hbar^2 k^2}{2m_r^*}, \quad (11.15)$$

где $E_g = E_c - E_v$ — ширина запрещенной зоны, а m_r^* — приведенная масса электрона дырки, определяемая соотношением

$$\frac{1}{m_r^*} = \frac{1}{m_n^*} + \frac{1}{m_p^*}. \quad (11.16)$$

Коэффициент поглощения $\alpha(\hbar v)$ согласно (11.6) обратно пропорционален длине свободного пробега фотона l_Φ , которая в свою очередь определяется скоростью движения фотона v_Φ и временем его свободного пробега τ_Φ . Следовательно,

$$\alpha(\hbar v) = \frac{1}{l_\Phi} = \frac{1}{v_\Phi \tau_\Phi} = g(\hbar v) \frac{\hbar}{c}. \quad (11.17)$$

Здесь $g(\hbar v) = 1/\tau_\Phi$ — вероятность поглощения фотона в единицу времени; $v_\Phi = c/\hbar$ — скорость движения фотона в веществе; $\hbar$ — показатель преломления вещества (действительная часть показателя); c — скорость света в вакууме.

Количество поглощенных в единицу времени квантов света с энергией, лежащей в интервале от $\hbar v$ до $\hbar(v + dv)$, пропорционально вероятности перехода электронов $P(\hbar v)$, числу квантовых состояний в валентной зоне в интервале энергии $|dE|$:

$$g(\hbar v) d(\hbar v) = 2P(\hbar v) N(E) |dE|. \quad (11.18)$$

Здесь коэффициент 2 учитывает два возможных направления поляризации света. Плотность квантовых состояний в валентной зоне согласно (4.11) имеет вид:

$$N(E) = \frac{(2m_p^*)^{3/2}}{2\pi^2 \hbar^3} (E_v - E)^{1/2}. \quad (11.19)$$

Используя (11.14) и (11.19), получаем:

$$g(\hbar v) d(\hbar v) = \frac{2}{\pi^2} P(\hbar v) k^2 dk. \quad (11.20)$$

Но из выражения (11.15) следует, что

$$d(\hbar v) = \frac{\hbar^2}{m_r^*} k dk, \quad (11.21)$$

а

$$k = \frac{(2m_r^*)^{1/2}}{\hbar} (\hbar v - E_g)^{1/2} \quad (11.22)$$

поэтому

$$g(h\nu) = \frac{(2m_r^*)^{3/2}}{\pi^2 \hbar^3} P(h\nu) (h\nu - E_g)^{1/2}. \quad (11.23)$$

Для прямых разрешенных переходов с сохранением квазинимпульса (11.11) вероятность перехода $P(h\nu)$ не зависит от энергии фотона. С учетом этого коэффициента собственного поглощения для прямых разрешенных переходов (см. рис. 11.2) на основании (11.17) и (11.23) выражается формулой

$$\left. \begin{aligned} \alpha &= A (h\nu - E_g)^{1/2}, & h\nu > E_g; \\ \alpha &= 0, & h\nu \leq E_g. \end{aligned} \right\} \quad (11.24)$$

где

$$A = \frac{\pi e^2 [2m_n^* m_p^* / (m_n^* + m_p^*)]^{3/2}}{\hbar^2 c m_n^* e_r \bar{n}}. \quad (11.25)$$

Соотношение (11.24) справедливо в ограниченном интервале изменения $(h\nu - E_g)$. Как следует из рис. 11.4, α^2 линейно зависит от $h\nu$ в некотором интервале. Продолжение этой прямой до пересечения с осью абсцисс позволяет определить ширину запрещенной зоны E_g для прямых разрешенных переходов.

Из (11.24) следует, что для прямых переходов не должно иметь место поглощение квантов с энергией, меньшей ширины запрещенной зоны. В силу этого край собственного поглощения со стороны длинных волн (или малых энергий) должен быть очень резким. Действительно, как следует из рис. 11.5, у очень чистого монокристалла антимида индия, имеющего прямые зоны, наблюдается крутой рост поглощения.

В случае образования валентной зоны из s -состояния, а зоны проводимости на основе атомной p -функции переходы будут прямыми, но «запрещенными» (см. рис. 2.11). При запрещенных переходах для всех $k \neq 0$ вероятность перехода $P(h\nu)$ пропорциональна k^2 .



Рис. 11.4. Зависимость квадрата коэффициента поглощения при прямых разрешенных переходах от энергии

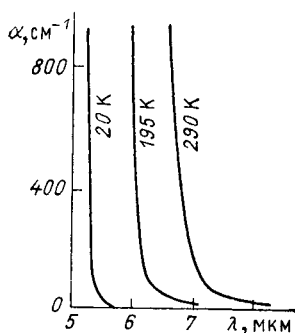


Рис. 11.5. Спектр поглощения антимида сурьмы при разных температурах

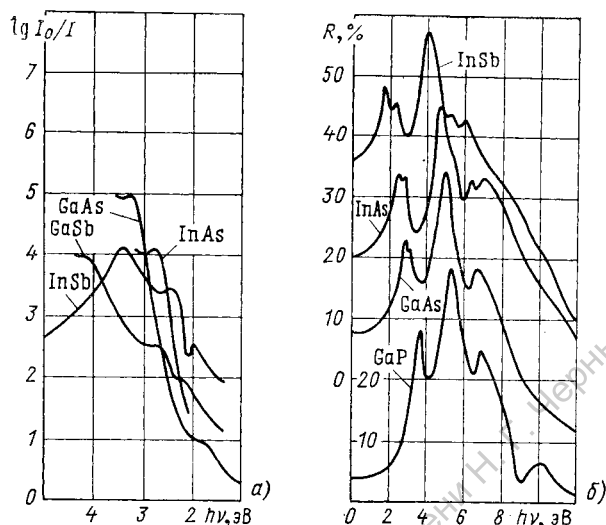


Рис. 11.6. Спектр поглощения при 80 К (а) и спектры отражения при 300 К (б) некоторых полупроводников типа $A^{III}B^V$

В силу этого для запрещенных прямых переходов вероятность перехода зависит от энергии как

$$P(h\nu) \approx k^2 = \text{const} (h\nu - E_g). \quad (11.26)$$

Поэтому коэффициент поглощения для прямых запрещенных межзонных переходов определяется соотношением вида

$$\alpha'(h\nu) = A' (h\nu - E_g)^{3/2}. \quad (11.27)$$

Зона проводимости полупроводников, как правило, включает несколько подзон (см. рис. 2.21). Поэтому при поглощении фотона с энергией, большей ширины запрещенной зоны, прямые переходы из валентной зоны в зону проводимости могут происходить во всех точках пространства квазиимпульса, для которых выполняется правило отбора. В этом случае спектр поглощения будет сложным, отражая сложную структуру зоны проводимости. Для таких исследований используются тонкие пленки полупроводников, поскольку при энергиях фотона $E > E_g$ коэффициент поглощения α очень велик. В качестве примера на рис. 11.6, а представлены спектры поглощения тонких пленок некоторых полупроводников типа $A^{III}B^V$ при 80 К.

Прямые переходы при больших энергиях будут проявляться и в спектрах отражения. В них, как видно из рис. 11.6, б, наблюдается серия пиков, отражающая сложное строение зоны проводимости. Большая интенсивность этих пиков обусловлена высокой плотностью начальных и конечных состояний.

11.3. СОБСТВЕННОЕ ПОГЛОЩЕНИЕ ПРИ НЕПРЯМЫХ ПЕРЕХОДАХ

В некоторых случаях существуют факторы, «смягчающие» правила отбора, в результате чего оказываются допустимыми также не-вертикальные переходы между непрямыми долинами (переход 2 на рис. 11.2, б). Однако вероятность таких переходов намного меньше вероятности прямых переходов (переход 1 на рис. 11.2, б). В отличие от прямых переходов закон сохранения квазиимпульса при не-пря-мых переходах обеспечивается взаимодействием с фононом, как это пока-зано на рис. 11.7, а. Следовательно, не-прямые, или не-вертикальные, переходы осуществляются с испусканием или поглощением фононов. Эти переходы определяют поглощение, расположенное с длинноволно-вой стороны у границы собственного поглощения, обусловленного пря-мыми переходами. Прямые и не-прямые переходы достаточно четко проявляются в спектре поглощения германия и кремния. Согласно рис. 11.8 пороговые значения энергии фононов в германии, соответ-ствующие не-прямым и прямым оптическим переходам, при комнат-ной температуре равны 0,64 и 0,81 эВ.

У полупроводника с непрямыми долинами в случае не-прямых переходов возможны переходы из любого занятого состояния валент-ной зоны в любое свободное состояние зоны проводимости. При рас-чете коэффициента поглощения для не-прямых переходов между не-прямыми долинами, которые сопровождаются поглощением или испу-сканием одного фонона, можно поступить следующим образом. Непрямой переход будем рассматривать как переход, происходящий через ряд промежуточных «виртуальных» состояний, имеющих очень малое время жизни. В этом случае выполнение закона сохра-нения энергии имеет место только для всего перехода в целом, в то время как при переходе электрона в виртуальное состояние его ква-зиимпульс сохраняется.

Для простоты положим, что не-прямой переход между непрямыми долинами происходит из начального состояния $k_1 = 0$ в конечное с $k_2 = k_{min}$ (переход 2' на рис. 11.7, б). Для такого перехода, идущего с испусканием или поглощением фонона, имеется две возможности. Электрон, находящийся вблизи потолка валентной зоны с $k_1 = 0$, возбуждаясь светом, переходит в зону проводимости, занимая там состояние с тем же волновым вектором (переход 1 на рис. 11.7, б). В валентной зоне в результате такого перехода остается дырка

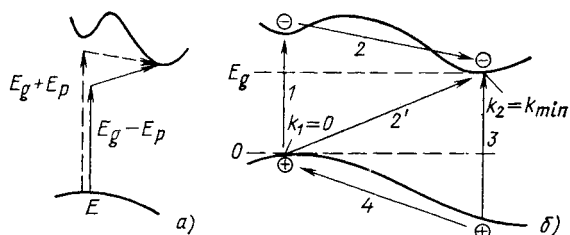


Рис. 11.7. Непрямые переходы в полупроводнике с непрямыми долинами

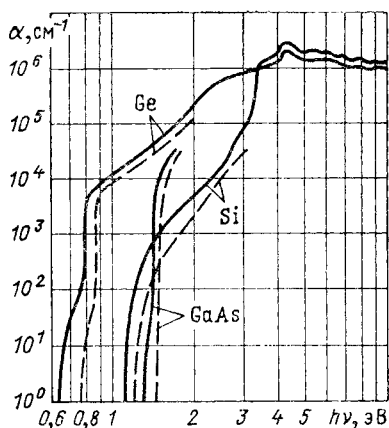


Рис. 11.8. Спектр поглощения чистых Ge, Si, GaAs

с $k_1 = 0$. Однако электрон проводимости с $k_1 = 0$ обладает большей энергией, чем та, которая соответствует дну зоны проводимости, поэтому он через очень малый промежуток времени совершит переход в состояние вблизи минимума энергии $k_2 = k_{min}$, поглощая или испуская при этом фотон (переход 2 на рис. 11.7, б).

Кроме того, возможен и такой переход. Поглощая квант света, электрон из глубоко лежащего состояния в валентной зоне вертикально переходит в состояние зоны проводимости с $k_2 = k_{min}$ (переход 3 на рис. 11.7, б). В результате в валентной зоне остается глубоко лежащая дырка, которая совершит переход в состояние

$k_1 = 0$ вблизи потолка валентной зоны, испуская или поглощая при этом фотон (переход 4 на рис. 11.7, б).

Произведем расчет коэффициента поглощения для переходов, когда вначале электрон совершает вертикальный переход 1, а затем осуществляет переход 2. Пусть начальная энергия электрона в валентной зоне E , а энергия электронов в конечном состоянии в зоне проводимости E' . Энергию фотона, принимающего участие в оптических переходах, обозначим через E_p . Если начало отсчета энергии производится от дна валентной зоны, то энергия кванта света, необходимая электрону для совершения непрямого перехода (рис. 11.7, б) в случае поглощения фотона, будет определяться равенством

$$h\nu_{\alpha} = E' - E - E_p, \quad (11.28)$$

а в случае испускания фотона

$$h\nu_e = E' - E + E_p. \quad (11.29)$$

Коэффициент поглощения α пропорционален интегралу по всем возможным парам состояний, разделенных энергией $h\nu \pm E_p$, от произведения плотностей начальных $N(E)$ и конечных состояний $N(E')$, а также вероятности взаимодействия с фотонами P , которая есть функция числа фононов N_p с энергией E_p . Следовательно,

$$\alpha = \int_0^{E_m} N_p N(E) N(E') dE. \quad (11.30)$$

С учетом принятого отсчета энергии (рис. 11.7) плотность начальных состояний при энергии E (в валентной зоне) имеет вид:

$$N(E) = \frac{(2m_p^*)^{3/2}}{2\pi^2 \hbar^3} |E|^{1/2}, \quad (11.31)$$

а плотность состояний при энергии E' (в зоне проводимости) на основе (11.28) и (11.29) равна

$$N(E') = \frac{(2m_n^*)^{3/2}}{2\pi^2\hbar^3} (E' - E_g)^{1/2} = \frac{(2m_n^*)^{3/2}}{2\pi^2\hbar^3} (\hbar\nu - E_g \pm E_p + E)^{1/2}. \quad (11.32)$$

Если неравновесная функция распределения фононов N_p мало отличается от равновесной $\langle n \rangle$, то согласно (3.75) можно считать, что число фононов определяется формулой

$$N_p = \langle n \rangle = \frac{1}{e^{E_p/kT} - 1}. \quad (11.33)$$

Так как $E_m = -(\hbar\nu - E_g \pm E_p)$, то

$$\begin{aligned} \alpha(\hbar\nu) &= A N_p \int_0^{-(\hbar\nu - E_g \pm E_p)} |E|^{1/2} (\hbar\nu - E_g \pm E_p + E)^{1/2} dE = \\ &= A N_p (\hbar\nu - E_g \pm E_p)^2, \end{aligned} \quad (11.34)$$

где A — некоторая постоянная.

Если подставить (11.33) в (11.34), то коэффициент поглощения для переходов с поглощением фонона приобретает вид:

$$\begin{aligned} \alpha_a(\hbar\nu) &= \frac{A(\hbar\nu - E_g - E_p)^2}{e^{E_p/kT} - 1}, \quad \hbar\nu > E_g - E_p; \\ \alpha_a &= 0, \quad \hbar\nu \leq E_g + E_p. \end{aligned} \quad (11.35)$$

Вероятность испускания фонона пропорциональна $N_p + 1$, поэтому коэффициент поглощения с испусканием фонона равен:

$$\alpha_e(\hbar\nu) = \frac{A(\hbar\nu - E_g + E_p)^2}{1 - e^{-E_p/kT}}, \quad \hbar\nu > E_g + E_p. \quad (11.36)$$

Поскольку при $\hbar\nu > (E_g + E_p)$ возможно поглощение света как с поглощением, так и с испусканием фононов, то коэффициент поглощения при не прямых межзонных переходах будет иметь вид суммы:

$$\begin{aligned} \alpha(\hbar\nu) &= \alpha_a(\hbar\nu) + \alpha_e(\hbar\nu) = \\ &= A \left[\frac{(\hbar\nu - E_g + E_p)^2}{e^{E_p/kT} - 1} + \frac{(\hbar\nu - E_g - E_p)^2}{1 - e^{-E_p/kT}} \right], \quad \hbar\nu > E_g + E_p. \end{aligned} \quad (11.37)$$

Соотношение (11.37) можно представить в виде

$$\begin{aligned} \alpha &= A \frac{(\hbar\nu - E_g + E_p)^2}{e^{E_p/kT} - 1} \quad \text{для } E_g - E_p < \hbar\nu \leq E_g + E_p; \\ \alpha &= 0 \quad \text{для } \hbar\nu \leq E_g - E_p. \end{aligned} \quad (11.38)$$

График зависимости $\alpha^{1/2}$ от $\hbar\nu$ для германия и кремния, как следует из рис. 11.9, имеет два прямолинейных участка. Прямая 1 соот-

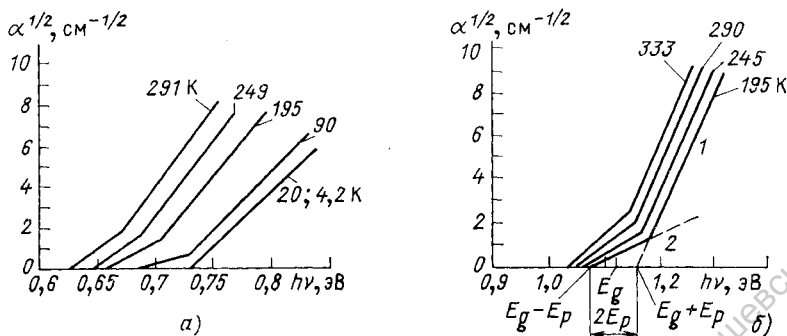


Рис. 11.9. Зависимость $\alpha^{1/2}$ от температуры для германия (а) и кремния (б)

ветствует зависимости $\alpha_e^{1/2}$, которая пересекается с осью абсцисс в точке $E_g + E_p$. Прямая 2 характеризует зависимость $\alpha_a^{1/2}$ и пересекается с осью абсцисс в точке $E_g - E_p$. Длина отрезка между точками пересечения прямых $\alpha_a^{1/2}$ и $\alpha_e^{1/2}$ с осью $h\nu$, соответствующих одной и той же температуре, равна $2E_p$. Посередине этого отрезка находится точка $h\nu = E_g$. При этом, как следует из рис. 11.9, при понижении температуры исследования график зависимости $\alpha^{1/2}$ смещается в сторону больших значений $h\nu$, что отражает температурную зависимость ширины запрещенной зоны, а наклон прямой $\alpha_a^{1/2}$ стремится к нулю. Это происходит в силу того, что при низких температурах возбуждается очень мало фононов, а поэтому невозможны оптические переходы с поглощением фононов.

Если прямой переход при $k_1 = 0$ запрещен, то непрямые запрещенные межзонные переходы для этого случая будут характеризоваться коэффициентом поглощения, который согласно (11.26) будет пропорционален

$$\alpha' \sim \frac{(h\nu - E_g + E_p)^3}{e^{E_p/kT} - 1} \quad \text{для } E_g - E_p < h\nu \leq E_g + E_p. \quad (11.39)$$

Таблица 11.1. Энергии фононов и дебаевские температуры для различных ветвей колебаний в германии и кремнии

Полупроводник	Тип колебаний	Энергия фононов $E_p = k\theta$, эВ	Эквивалентная температура θ , К
Германий	Продольные акустические LA	0,0275	320
	Поперечные акустические TA	0,0077	90
	Поперечные оптические TO	0,034	400
Кремний	Продольные акустические LA	0,016	230
	Поперечные акустические TA	0,055	700
	Поперечные оптические TO	0,119	1370
	Продольные оптические LO	0,083	1010

Для прямых разрешенных межзонных переходов коэффициент поглощения α обычно достигает 10^4 — 10^5 см $^{-1}$, а для непрямых переходов между непрямыми долинами $\alpha = 10^{-1} \div 10^3$ см $^{-1}$ (см. рис. 11.8).

При рассмотрении непрямых межзонных переходов учитывался только один тип фононов. В действительности в кристалле может существовать много различных фононов. В оптических переходах принимают участие только те фононы, которые обеспечивают необходимое изменение квазиимпульса электрона. Такими фононами обычно оказываются продольные и поперечные акустические фононы. В табл. 11.1 приведены для германия и кремния значения энергии фононов E_p , а также дана эквивалентная им температура $\theta = E_p/k$.

11.4. ПОГЛОЩЕНИЕ СИЛЬНО ЛЕГИРОВАННОГО И АМОРФНОГО ПОЛУПРОВОДНИКОВ

У антимонида индия, имеющего прямую запрещенную зону, по мере роста концентрации свободных электронов край полосы собственного поглощения смещается в сторону коротких длин волн. Например, у материала n -типа при концентрации электронов 5×10^{18} см $^{-3}$ край поглощения приходится на 3,5 мкм, а у легированного полупроводника длинноволновая граница поглощения при комнатной температуре соответствует 7,1 мкм (рис. 11.10, а). Это обусловлено тем, что у антимонида индия из-за малой эффективной массы электронов ($m_n^* = 0,013m_0$) зона проводимости характеризуется большой кривизной $\frac{1}{m^*} = \frac{1}{\hbar} \frac{\partial^2 E}{\partial k^2}$, поэтому плотность состояний в ней мала (рис. 11.10, б). Вследствие этого при увеличении концентрации свободных электронов быстро наступает вырождение электронного газа и уровень Ферми F_n поднимается выше дна зоны проводимости на $\xi_n = F_n - E_c$. Поскольку состояния, расположенные ниже уровня Ферми, уже заполнены, поглощение, связанное с переходами в эти состояния, невозможно. В этом случае $h\nu_{min}$ (прямой переход на рис. 11.10, б) будет больше ширины запрещенной зоны, а поэтому край поглощения должен быть смещен в сторону больших энергий на величину, примерно равную ξ_n .

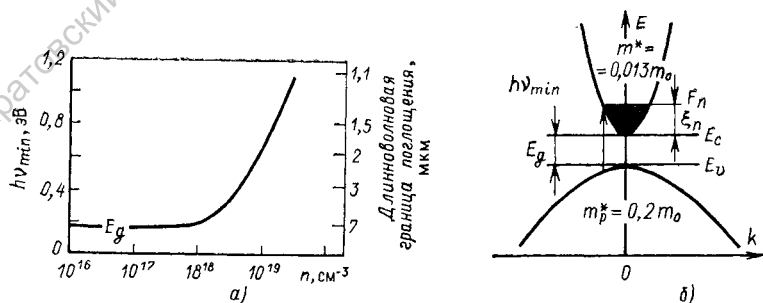


Рис. 11.10. Зависимость оптического энергетического зазора $h\nu_{min}$ от концентрации электронов (а) и энергии от квазиимпульса (б) в InSb n -типа

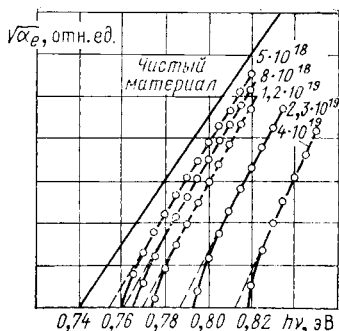


Рис. 11.11. Край поглощения германия, сильно легированного мышьяком, при 4,2 К

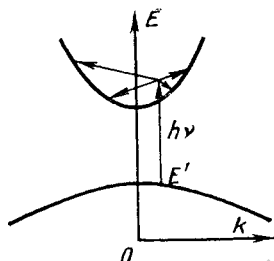


Рис. 11.12. Четыре из возможных не прямых переходов из начального состояния E' в зону проводимости полупроводника с прямой запрещенной зоной

В сильно легированном полупроводнике с непрямыми долинами также возможно смещение края поглощения, связанное с заполнением зон (рис. 11.11). В этом случае сохранение квазиимпульса электрона при не прямых переходах обеспечивается за счет электрон-электронного взаимодействия или за счет рассеяния на примеси. При этом вероятность рассеяния пропорциональна числу рассеивающих центров N , и поскольку участия фононов не требуется, то коэффициент поглощения описывается формулой

$$\alpha = AN(h\nu - E_g - \xi_n)^2, \quad (11.40)$$

где $\xi_n = F_n - E_c$ — положение уровня Ферми в зоне проводимости сильно легированного полупроводника.

У полупроводников с прямыми долинами возможны и не прямые переходы (рис. 11.12). Такие переходы, будучи процессами второго порядка, обладают меньшей вероятностью, чем прямые переходы, поэтому они обуславливают незначительное поглощение, вызывая удлинение края полосы собственного поглощения (см. рис. 11.5). Эти не прямые переходы похожи на не прямые переходы между не прямыми долинами (см. рис. 11.7). Для них сохранение квазиимпульса электрона также обеспечивается за счет поглощения или излучения фонона, а также за счет рассеяния на свободных носителях заряда или на примесях. Здесь также возможны переходы из любого занятого состояния валентной зоны в любое свободное состояние зоны проводимости. Поэтому коэффициент поглощения в этом случае определяется формулами (11.35) — (11.38), если закон сохранения квазиимпульса обеспечивается участием фононов, или формулой (11.40), если в акте поглощения фононы не принимают участия.

Кроме того, у некоторых сильно легированных полупроводников с прямыми долинами наблюдается поглощение вблизи края, которое возрастает по экспоненциальному закону. Как известно, увеличение концентрации примесей сопровождается образованием вблизи раз-

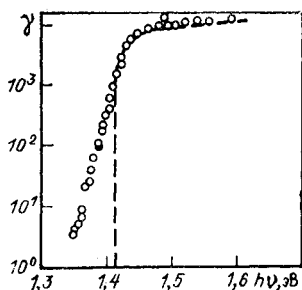


Рис. 11.13. Край поглощения GaAs при комнатной температуре: сплошные кружки — экспериментальные данные; пунктир — теоретическая кривая $\alpha = A(h\nu - E_g)^{1/2}$

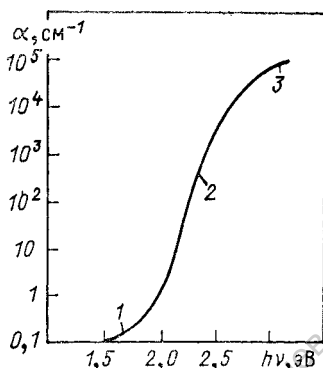


Рис. 11.14. Зависимость коэффициента поглощения от энергии в стеклообразном полупроводнике

решенных зон хвостов плотности состояний в пределах запрещенной зоны (см. § 4.9). Переходы между хвостами зон, форма и величина которых зависят от легирования материала, определяют, в частности, экспоненциальный край поглощения в арсениде галлия (рис. 11.13).

На рис. 11.14 представлен спектр поглощения стеклообразного полупроводника As_2S_3 . Эту спектральную зависимость можно условно разделить на три части:

1 — длинноволновый участок поглощения, который, возможно, определяется примесями или дефектами, его форма, как правило, иногда описывается правилом Урбаха:

$$\alpha = A \frac{a(h\nu - E_0)}{kT}, \quad (11.41)$$

где A , a , E_0 — константы;

2 — область поглощения, описываемая экспоненциальной зависимостью коэффициента поглощения от энергии, обусловлена переходами из расширенных состояний вблизи вершины валентной зоны на локализованные состояния вблизи дна зоны проводимости и переходами из локализованных состояний вблизи вершины валентной зоны на расширенные состояния в зоне проводимости;

3 — область сильного поглощения, которая обусловлена переходами из расширенных состояний возле вершины валентной зоны на расширенные состояния в зоне проводимости, и описывается степенным законом вида:

$$\alpha = \frac{(h\nu - E_g^0)^r}{h\nu}, \quad (11.41a)$$

где E_g^0 — «оптическая» запрещенная зона, которая равна разности E_c и E_v (рис. 4.22, а); $r = 2$.

11.5. ВЛИЯНИЕ ВНЕШНИХ ВОЗДЕЙСТВИЙ НА СОБСТВЕННОЕ ПОГЛОЩЕНИЕ ПОЛУПРОВОДНИКОВ

Многочисленные опыты, проведенные на разных полупроводниках, показывают, что внешние воздействия, такие как температура, давление, электрические и магнитные поля, оказывают существенное влияние на их спектр собственного поглощения. Знание влияния их на энергетические уровни полупроводника необходимо как для решения теоретических задач, связанных с рассмотрением поведения электронов в кристалле, так и для решения задач практического значения, связанных с изготовлением и эксплуатацией полупроводниковых приборов.

1. Влияние температуры

У полупроводника с повышением температуры происходят усиление колебаний атомов относительно положения равновесия и расширение решетки. У некоторых полупроводников согласно рис. 2.7 с увеличением постоянной решетки имеет место уменьшение ширины запрещенной зоны, которое сопровождается смещением полосы фундаментального поглощения в сторону длинных волн (см. рис. 11.5). Для многих полупроводников в области температур много меньших температуры Дебая, температурная зависимость ширины запрещенной зоны описывается эмпирическим соотношением вида

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta}, \quad (11.42)$$

где $E_g(0)$ — ширина запрещенной зоны при 0 К; α и β — константы, значения которых для некоторых полупроводников приведены в табл. 11.2.

Т а б л и ц а 11.2. Значение E_g и параметров α , β

Полупроводник	Тип межзонных переходов ¹	$E_g(0)$, эВ	$\alpha \cdot 10^4$	β
Si	E_{gi}	1,1557	7,021	1108
Ge	E_{gi}	0,7412	4,561	210
Ge	E_{gd}	0,8893	6,842	398
6H-SiC	E_{gi}	3,024	-0,3055	-311
GaAs	E_{gd}	1,5216	8,871	572
InP	E_{gd}	1,4206	4,906	327
InAs	E_{gd}	0,426	3,158	93

¹ Здесь E_{gi} — не прямые переходы; E_{gd} — прямые переходы.

При температурах, много больших температуры Дебая, зависимость ширины запрещенной зоны становится линейной:

$$E_g(T) = E_g(0) - \alpha T. \quad (11.43)$$

2. Влияние давления

При всестороннем сжатии кристалла, обеспечиваемом так называемым гидростатическим давлением, расстояние между его атомами уменьшается. Но уменьшение межатомного расстояния согласно рис. 2.7 должно вызвать увеличение ширины запрещенной зоны.

Увеличение E_g с ростом давления наблюдается у большинства полупроводников с прямыми запрещенными зонами. У таких полупроводников с ростом гидростатического давления край собственного поглощения смещается в сторону коротких длин волн.

У полупроводников с непрямыми запрещенными зонами может быть сложная зависимость E_g от давления. Например, у германия, как следует из рис. 11.15, с увеличением давления E_g возрастает, а у кремния, наоборот, уменьшается (рис. 11.16). У арсенида галлия с ростом гидростатического давления минимум зоны проводимости для прямых переходов движется вверх быстрее, чем минимум для непрямых переходов. Поэтому при повышении давления он может стать полупроводником с непрямыми переходами.

Одностороннее сжатие или растяжение кристалла приводит к сдвигу положения различных минимумов энергии в зоне проводимости. Например, сжатие кристалла кремния вдоль оси $[100]$ уменьшает ширину запрещенной зоны в направлении сжатия, но из-за одновременного растяжения в поперечном направлении минимумы зоны проводимости, находящиеся в плоскости, перпендикулярной направлению сжатия, смещаются в область больших энергий (рис. 11.16). В случае германия n -типа наибольший эффект наблюдается при сжатии в направлении $[110]$. Изменение энергетической структуры кристалла вызовет перераспределение электронов, которые переходят в состояния с меньшей энергией, возникшие в результате деформации сдвига. Это скажется на значении упругих констант, т. е. на коэффициенте пьезосопротивления.

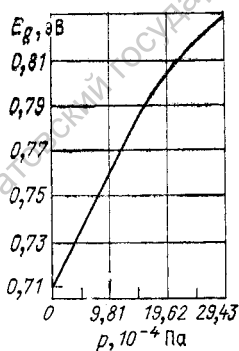


Рис. 11.15. Зависимость ширины запрещенной зоны германия при 350 К от давления

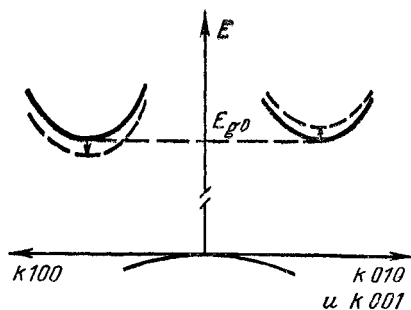


Рис. 11.16. Продольные и поперечные смещения долин зоны проводимости в кремнии под действием одноосного сжатия вдоль направления $[100]$

3. Влияние сильного электрического поля

Как было установлено (§ 2.10), во внешнем электрическом поле у полупроводника происходит наклон энергетических зон. В этом случае электроны валентной зоны могут туннелировать в зону проводимости.

Согласно рис. 11.17, *а* высота треугольного барьера при наклоне энергетических зон равна E_g , а его толщина d на основании (6.127)

$$d = E_g / e\mathcal{E}. \quad (11.44)$$

Как видно, с увеличением напряженности внешнего электрического поля $\mathcal{E}$ толщина барьера уменьшается, а следовательно, увеличивается вероятность процесса туннелирования электронов.

В присутствии внешнего электрического поля, как следует из рис. 11.17, *б*, имеет место поглощение квантов света с энергией $h\nu$, меньшей, чем ширина запрещенной зоны полупроводника, что эквивалентно уменьшению толщины барьера до значения d' . В этом случае

$$d' = (E_g - h\nu) / e\mathcal{E} \quad (11.45)$$

и туннельный переход становится более вероятным.

Процесс туннелирования электронов в присутствии внешнего электрического поля, идущий с поглощением или фотонов, принято называть эффектом Франца — Келдыша. В соответствующем полупроводнике он проявляется как сдвиг края поглощения в сторону меньших энергий (рис. 11.18).

4. Влияние магнитного поля

Рассмотрим вначале влияние магнитного поля на энергетический спектр кристалла. Эта задача может быть решена на основе уравнения Шредингера, в котором гамильтониан должен определять поведение электрона в магнитном поле.

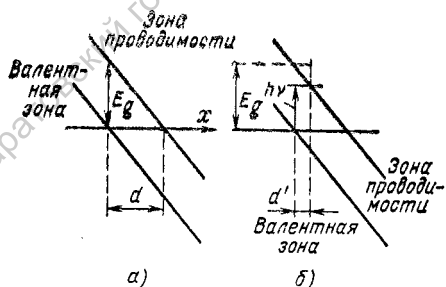


Рис. 11.17. Туннелирование электронов:
а — без изменения энергии; *б* — с поглощением фотона

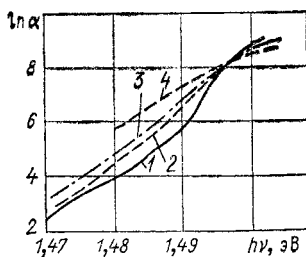


Рис. 11.18. Край поглощения GaAs при разной напряженности электрического поля:

1 — нулевая; 2 — 6 кВ/см; 3 — 10 кВ/см; 4 — 30 кВ/см

Из теории электромагнитного поля известно, что вектор индукции магнитного поля $\mathbf{B}$ связан с векторным потенциалом $\mathbf{A}(x, y, z)$ соотношением

$$\mathbf{B} = \text{rot } \mathbf{A}. \quad (11.46)$$

При этом электрон с зарядом $-e$, имеющий импульс $\mathbf{p}$, при прохождении в области действия поля, определенного векторным потенциалом $\mathbf{A}(x, y, z)$, будет обладать обобщенным импульсом

$$\mathbf{P} = \mathbf{p} + e\mathbf{A}. \quad (11.47)$$

В этом случае, поскольку кинетическая энергия $T = \frac{p^2}{2m_0}$, функция Гамильтона для электрона запишется в виде

$$H = T + U(x, y, z) = \frac{1}{2m_0}(\mathbf{p} + e\mathbf{A})^2 + U(x, y, z), \quad (11.48)$$

где $U(x, y, z)$ — потенциальная энергия.

Положим для определенности, что однородное магнитное поле направлено по оси z , т. е.

$$B_x = B_y = 0; \quad B_z = B. \quad (11.49)$$

Так как

$$\begin{aligned} \text{rot } \mathbf{A} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = \mathbf{i} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \mathbf{j} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \\ + \mathbf{k} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right), \end{aligned}$$

то на основании (11.49) имеем:

$$\left. \begin{aligned} B_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = 0; \\ B_y &= \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} = 0; \\ B_z &= \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} = B. \end{aligned} \right\} \quad (11.50)$$

Уравнения (11.50) удовлетворяются, если положить, например,

$$A_x = A_z = 0; \quad A_y = Bx. \quad (11.51)$$

С учетом (11.51) функция Гамильтона (11.48) определится соотношением

$$H = \frac{1}{2m_0} [p_x^2 + (p_y^2 + eA_y)^2 + p_z^2] + U(x, y, z). \quad (11.52)$$

Воспользовавшись приближением эффективной массы (см. § 2.14), запишем уравнение Шредингера $\hat{H}\Psi = E\Psi$, переходя от функции

Гамильтона к оператору Гамильтона с заменой в (11.52) p_x на $\frac{\hbar}{i} \times \times \frac{\partial}{\partial x}$, p_y на $\frac{\hbar}{i} \frac{\partial}{\partial y}$ и p_z на $\frac{\hbar}{i} \frac{\partial}{\partial z}$. В результате получим:

$$-\frac{\hbar^2}{2m^*} \left[\frac{\partial^2 \Psi}{\partial x^2} + \left(\frac{\partial}{\partial y} + \frac{ieBx}{\hbar} \right)^2 \Psi + \frac{\partial^2 \Psi}{\partial z^2} \right] = E\Psi. \quad (11.53)$$

В (11.53) z и y в явном виде не содержатся. К тому же в отсутствие магнитного поля (11.53) переходит в известное уравнение для волн де Бройля — $\frac{\hbar^2}{2m^*} \Delta \Psi_0 = E\Psi_0$, которое имеет решения:

$$\Psi_0 = C e^{i(k_x x + k_y y + k_z z)} \text{ и } E_0 = \frac{\hbar^2}{2m^*} (k_x^2 + k_y^2 + k_z^2). \quad (11.54)$$

Отсюда решение уравнения (11.52) можно искать в виде суперпозиции

$$\Psi(x, y, z) = \varphi(x) e^{i(k_y y + k_z z)}, \quad (11.55)$$

где $\varphi(x)$ — некоторая функция, зависящая только от x .

Подставив (11.55) в уравнение Шредингера (11.53) и сократив на экспоненту, получим:

$$-\frac{\hbar^2}{2m^*} \frac{d^2 \varphi}{dx^2} + \frac{\hbar^2 k_y^2}{2m^*} \varphi + \frac{\hbar e k_y B x}{m^*} \varphi + \frac{e^2 B^2 x^2}{2m^*} \varphi + \frac{\hbar^2 k_z^2}{2m^*} \varphi = E\varphi. \quad (11.56)$$

Введем обозначения

$$\omega_0 = \frac{eB}{m^*}; \quad x_0 = -\frac{\hbar}{eB} k_y; \quad E_B = E - \frac{\hbar^2 k_z^2}{2m^*} \quad (11.57)$$

и перепишем в них (11.56). В результате получим:

$$-\frac{\hbar^2}{2m^*} \frac{d^2 \varphi}{dx^2} + \frac{1}{2} m^* \omega_0^2 (x - x_0)^2 \varphi = E_B \varphi. \quad (11.58)$$

Уравнение (11.58) представляет собой уравнение гармонического осциллятора. Таким образом, движение электрона по направлению x описывается уравнением линейного гармонического осциллятора с массой m^* и собственной частотой ω_0 , совершающего колебания около положения равновесия x_0 . Но энергия гармонического осциллятора

$$E_B = \hbar \omega_0 \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (11.59)$$

и, следовательно, энергия электрона в магнитном поле

$$E = E_c + \frac{\hbar^2 k_z^2}{2m^*} + E_B = E_c + \frac{\hbar^2 k_z^2}{2m^*} + \hbar \omega_0 \left(n + \frac{1}{2} \right). \quad (11.60)$$

Отсюда следует, что если вектор индукции магнитного поля направлен по оси z , то поле не оказывает никакого влияния на движение электрона в этом направлении и энергия, связанная с этим движением, не квантована, т. е. электрон вдоль магнитного поля

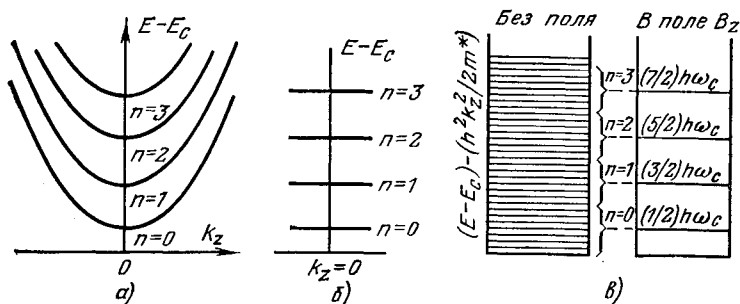


Рис. 11.19. Спектр энергии электронов кристалла в магнитном поле

движется со скоростью $v_z = \frac{\hbar(k_z - k_{0z})}{m_n^*}$. В плоскости xoy , перпендикулярной индукции $\mathbf{B}$, электрон совершает движение по круговой орбите с циклотронной частотой $\omega_0 = \omega_c = eB/m^*$ (см. § 2.13), причем для перехода его с одной круговой орбиты на другую требуется энергия $\hbar\omega_0$. Таким образом, энергия движения электрона в плоскости xoy под действием магнитного поля становится квантованной и спектр поглощенной энергии, связанной с переходами электрона между разрешенными орбитами, будет линейчатым.

Из (11.60) следует, что спектр энергии $E(k_z)$ электронов кристалла в каждой зоне представляет собой набор парабол, сдвинутых по оси энергии относительно друг друга на $\hbar\omega_0$ (рис. 11.19, а). Для $k_z = 0$ спектр энергии изображается совокупностью дискретных уровней (рис. 11.19, б), которые возникают в результате группировки квазинепрерывного спектра (рис. 11.19, в). Уровни энергии электрона кристалла, находящегося в магнитном поле, называют уровнями Ландау.

Магнитное поле не только приводит к квантованию энергии электронов кристалла, но и сдвигает дно зоны проводимости и дно валентной зоны. Действительно, согласно (11.60) минимальная кинетическая энергия электрона в зоне проводимости ($n = 0$) не равна нулю, а составляет $\frac{1}{2} \hbar\omega_0$, т. е. в магнитном поле дно зоны проводимости смещается вверх на величину

$$\frac{\hbar\omega_0}{2} = \frac{\hbar eB}{2m^*}. \quad (11.61)$$

Соответственно минимальная энергия дырок с эффективной массой m_p^* в магнитном поле возрастает на $\hbar eB/2m_p^*$, что выражается в смещении вниз на это значение потолка валентной зоны. Таким образом, под действием магнитного поля ширина запрещенной зоны полупроводника возрастает на величину ΔE_g , равную:

$$\Delta E_g = \frac{e\hbar B}{2} \left(\frac{1}{m_n^*} + \frac{1}{m_p^*} \right) = \frac{e\hbar B}{2m_e^*}, \quad (11.62)$$

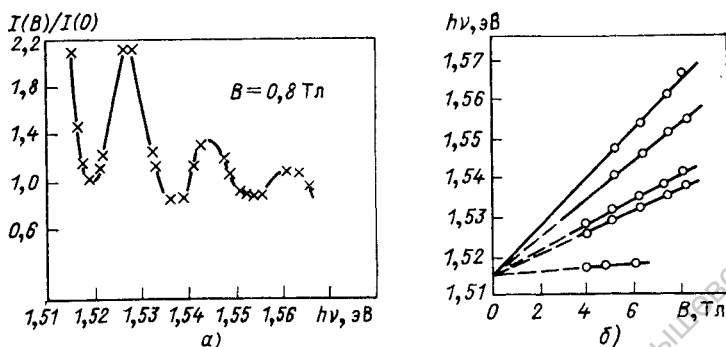


Рис. 11.20. Магнитно-осцилляционный эффект поглощения в арсениде галлия при 77 К

где $\frac{1}{m_r^*} = \frac{1}{m_n^*} + \frac{1}{m_p^*}$ — обратная величина приведенной эффективной массы.

Из равенства (11.61) следует, что сдвиг края соответствующей зоны в магнитном поле обратно пропорционален эффективной массе носителей заряда в ней. Для антимонида индия $m_n^* = 0,013m_0$ сдвиг края зоны проводимости составляет $4,85 \cdot 10^{-5}$ (эВ·м)/А.

Увеличение ширины запрещенной зоны полупроводника, помещенного в магнитное поле, выразится в сдвиге края полосы фундаментального поглощения в коротковолновую сторону, причем сдвиг пропорционален напряженности магнитного поля. Экспериментально, измеряя его величину, на основании (11.62) легко рассчитать приведенную эффективную массу носителей заряда.

Появление уровней Ландау при помещении полупроводника в магнитное поле создает возможность для переходов электронов между этими уровнями при поглощении квантов света с энергией, кратной $\hbar\omega_c$. Это должно привести к появлению линейчатой структуры на спектральной кривой поглощения. Однако поскольку энергия электрона при движении по направлению вектора напряженности магнитного поля не квантована, то наблюдается некоторое размазывание линейчатого спектра в осциллирующий с последовательностью узких полос с максимумами и минимумами (рис. 11.20, а). Этот эффект получил название магнитно-осцилляционного эффекта поглощения или магнетопоглощения.

Магнетопоглощение легче наблюдается в случае прямых переходов, так как тонкая осцилляторная структура при непрямых переходах может оказаться затухающей. Кроме того, для экспериментального наблюдения магнетопоглощения должно выполняться очевидное условие

$$\hbar\omega_c \gg kT, \quad (11.63)$$

т. е. расстояние между уровнями энергии в магнитном поле должно быть больше тепловой энергии электронов, иначе квантованный спектр электронов кристалла будет квазинепрерывным.

Из рис. 11.20, а следует, что расстояния между соседними максимумами или минимумами на спектральной зависимости магнетопоглощения составляют $\hbar\omega_0$. Поэтому если экспериментально снять спектр магнетопоглощения в магнитных полях и построить зависимость положения максимумов осциллирующего спектра от индукции магнитного поля, то эти зависимости будут представлять собой прямые, сходящиеся к одной точке в нулевом магнитном поле (рис. 11.20, б). Согласно соотношению (11.62) изменение ширины запрещенной зоны полупроводника в этой точке равно нулю и ее положение соответствует минимальной энергии, необходимой для осуществления прямого перехода между зонами, т. е. оптической ширины запрещенной зоны. Измерения, проведенные таким методом, дали ширину запрещенной зоны при 300 К в точке $k = 0$ 0,803 эВ для германия, 1,12 эВ для кремния и 1,43 эВ для арсенида галлия. Эти значения были получены для материалов достаточной степени чистоты.

Необходимо отметить, что расщеплением разрешенных энергетических зон на уровни Ландау объясняется как эффект магнетопоглощения, так и эффект циклотронного резонанса. Однако поскольку эффект магнетопоглощения обусловлен переходами между уровнями Ландау, образовавшимися в зоне проводимости и в валентной зоне, то спектр магнетопоглощения лежит внутри полосы фундаментального поглощения. Осциллирующий спектр циклотронного резонанса, обусловленный переходами между уровнями Ландау одной и той же зоны, расположен в области поглощения свободными носителями заряда.

11.6. ЭКСИТОННОЕ ПОГЛОЩЕНИЕ

При поглощении света полупроводником возможно такое возбуждение электрона валентной зоны, при котором он не переходит в зону проводимости, а образует с дыркой связанную квазичастицу, обладающую нулевым спином, поскольку электрон и дырка имеют антипараллельные спины. Такую систему, получившую название экситона, можно представить как возбужденное состояние атома кристалла, которое передается от атома к атому посредством квантовомеханического резонанса.

Если размеры экситона велики по сравнению с постоянной решетки, то взаимодействие электрона и дырки можно представить как кулоновское взаимодействие двух точечных зарядов, ослабленное средой в ϵ_r раз. Пусть $\mathbf{r}_n$ и $\mathbf{r}_p$ — радиус-векторы, определяющие положения электрона и дырки, а m_n^* и m_p^* — их скалярные эффективные массы. Тогда уравнение Шредингера для взаимодействующей пары электрон — дырка имеет вид:

$$\left[-\frac{\hbar^2}{2m_n^*} \nabla_{r_n}^2 - \frac{\hbar^2}{2m_p^*} \nabla_{r_p}^2 - \frac{e^2}{4\pi\epsilon_r\epsilon_0 (r_n - r_p)} \right] \Psi(\mathbf{r}_n, \mathbf{r}_p) = W_{\text{экс}} \Psi(\mathbf{r}_n, \mathbf{r}_p), \quad (11.64)$$

где $\nabla_{r_n}^2$ и $\nabla_{r_p}^2$ — операторы Лапласа в координатах электрона и дырки; $W_{\text{экс}}$ — энергия экситона.

Если ввести радиус-вектор центра тяжести R , характеризующий положение электронно-дырочной пары, и радиус-вектор r относительно положения электрона и дырки

$$R = \frac{m_n^* r_n + m_p^* r_p}{m_n^* + m_p^*}; \quad (11.65)$$

$$r = r_n - r_p, \quad (11.66)$$

то уравнение (11.64) в новых координатах будет иметь вид

$$\left[-\frac{\hbar^2}{2M^*} \nabla_R^2 - \frac{\hbar^2}{2m_r^*} \nabla_r^2 - \frac{e^2}{4\pi\epsilon_r\epsilon_0 |r|} \right] \Psi(R, r) = W_{\text{экс}} \Psi(R, r), \quad (11.67)$$

где

$$M^* = m_n^* + m_p^* \text{ и } m_r^* = \frac{m_n^* m_p^*}{m_n^* + m_p^*} \quad (11.68)$$

есть эффективная и приведенная масса экситона.

Собственные волновые функции экситонных состояний, удовлетворяющих уравнению (11.67), имеют форму

$$\Psi(R, r) = e^{i(k_{\text{экс}}R)} \varphi(r), \quad (11.69)$$

где $k_{\text{экс}} = k_n + k_p$ — волновой вектор экситона, а k_n и k_p — волновые векторы электрона и дырки. Плоская волна $e^{i(k_{\text{экс}}R)}$ описывает свободное движение экситона, а волновая функция $\varphi(r)$ описывает состояние внутреннего движения экситона, т. е. движение электрона и дырки относительно центра массы. Подстановка (11.69) в (11.67) сводит задачу к водородоподобной:

$$\left[-\frac{\hbar^2}{2m_r^*} \nabla_r^2 - \frac{e}{4\pi\epsilon_r\epsilon_0 |r|} \right] \varphi(r) = \left(W_{\text{экс}} - \frac{\hbar^2 k_{\text{экс}}^2}{2M^*} \right) \varphi(r) = E'_{\text{экс}} \varphi(r). \quad (11.70)$$

Собственные значения энергии $E'_{\text{экс}}$, являющиеся энергией связи экситона, образуют водородоподобную серию

$$E'_{\text{экс}} = -\frac{e^4 m_r^*}{32\pi^2 \hbar^2 \epsilon_r^2 \epsilon_0^2 n^2} = -\frac{E_{\text{экс}}}{n^2}, \quad (11.71)$$

где

$$E_{\text{экс}} = \frac{e^4 m_r^*}{32\pi^2 \hbar^2 \epsilon_r^2 \epsilon_0^2}; \quad (11.71a)$$

n — квантовое число, определяющее различные экситонные состояния ($n \geq 1$).

Поскольку в экситоне приведенная масса меньше эффективной массы электрона и дырки (11.68), энергия связи экситона будет меньше энергии ионизации донора (2.174).

Следовательно, полная энергия экситона с учетом (11.71) равна:

$$W_{\text{экс}} = \frac{\hbar^2 k_{\text{экс}}^2}{2M^*} + E_{\text{экс}} = \frac{\hbar^2 k_{\text{экс}}^2}{2M^*} - \frac{E_{\text{экс}}}{n^2}. \quad (11.72)$$

Примем дно зоны проводимости за начало отсчета энергии и будем считать, что ему соответствует начало сплошного спектра экситона ($n = \infty$). В этом случае энергетическая диаграмма различных экситонных состояний будет подобна представленной на рис. 11.21. Если оптический переход прямой совершается при $k = 0$, экситон будет характеризоваться водородоподобной серией линий поглощения, удовлетворяющих соотношению

$$h\nu = h\nu_{\infty} - \frac{E_{\text{экс}}}{n^2}, \quad (11.73)$$

где $h\nu_{\infty}$ совпадает с шириной запрещенной зоны при $k = 0$. Поэтому (11.73) можно переписать в виде

$$h\nu = E_g - \frac{E_{\text{экс}}}{n^2}, \quad (11.74)$$

где $E_{\text{экс}}$ — энергия связи экситона, определяемая зависимостью (11.71а).

В полупроводниках с прямыми переходами при $k = 0$ экситону ($n = 1$) соответствует очень узкий пик на краю поглощения, который уширяется с повышением температуры, как это показано на рис. 11.22.

Рис. 11.21. Диаграмма энергетических уровней экситона и его возбужденных состояний. Энергия экситона отсчитывается от края зоны проводимости

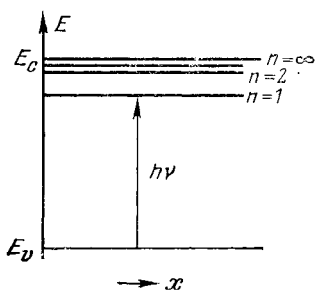
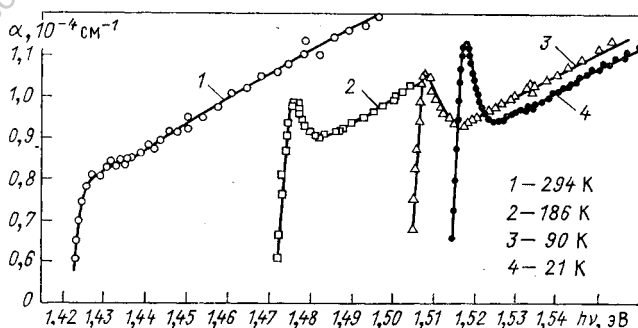


Рис. 11.22. Экситонное поглощение в арсениде галлия



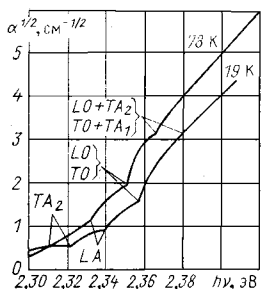
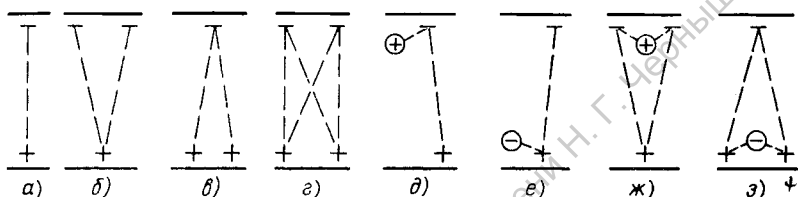


Рис. 11.23. Структура края поглощения GaP. Стрелками указаны пороговые энергии образования экситонов с эмиссией фононов

Рис. 11.24. Экситонные комплексы:

а, б, в, г — свободные экситоны; д, е, ж, з — связанные экситоны



В полупроводниках с непрямыми переходами также возможно образование экситонов. В этом случае длинноволновый край спектра поглощения экситона ($n = 1$) определяется из условия для переходов с поглощением фонона

$$h\nu = E_g - E_p - E_{\text{эк}} \quad (11.75)$$

и с испусканием фонона

$$h\nu = E_g + E_p - E_{\text{эк}}. \quad (11.76)$$

В полупроводниках с непрямыми переходами образование экситона сопровождается возникновением ступеньки на краю собственного поглощения. При этом возможны переходы с участием нескольких фононов, испускание и поглощение которых возможны в различных комбинациях, поэтому на краю поглощения будет наблюдаться много ступенек, как, например, на рис. 11.23.

При поглощении света полупроводником возможно и такое возбуждение электронов и дырок, когда кулоновское взаимодействие между ними приводит к объединению их с образованием ионоподобных или молекулоподобных комплексов. Так, два свободных электрона и одна свободная дырка или две свободные дырки и один свободный электрон, объединяясь, образуют экситонные ионы (рис. 11.24, б и в соответственно). Кроме того, два свободных электрона и две свободные дырки могут объединяться в комплекс (рис. 11.24, г), напоминающий молекулу позитрония (рис. 11.25).

Свободная дырка может объединиться с нейтральными донором или акцептором и образовать заряженные экситонные ионы (рис. 11.24, д, е). В этом случае электрон и дырка, связанные с донором, движутся вокруг донора (рис. 11.26), поэтому такой комплекс

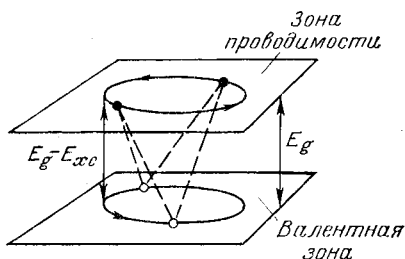


Рис. 11.25. Экситонный комплекс, состоящий из двух электронов, связанных с двумя дырками

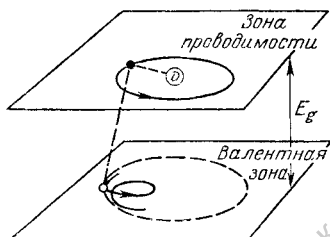


Рис. 11.26. Экситон, связанный с донором D

называется «связанным экситоном». Образование связанных экситонов может быть и в случае, когда два свободных электрона и одна свободная дырка связаны с донором (рис. 11.24, ж) или когда две дырки и электрон связаны с акцептором (рис. 11.24, з). Как свободные, так и связанные экситоны были обнаружены в ряде полупроводников.

11.7. ПОГЛОЩЕНИЕ СВОБОДНЫМИ НОСИТЕЛЯМИ ЗАРЯДА

При поглощении фотона электроны могут переходить в пределах той же зоны с одного уровня на другой (рис. 11.27). Такие внутризонные переходы носителей заряда происходят с нарушением правила отбора. В силу закона сохранения импульса они осуществляются либо когда наряду с поглощением фотона происходит поглощение или испускание фонона, либо когда имеется рассеяние на ионизованных примесях. Поэтому поглощение свободными электронами можно определить методами, которые используются для не прямых переходов. Однако результаты получаются те же, если определение коэффициента поглощения свободными электронами проводить на основе рассмотрения проводимости полупроводника в высокочастотном электрическом поле.

Как известно, теория распространения электромагнитных волн в проводящей среде основывается на уравнениях Максвелла, которые для случая, когда отсутствуют поля сторонних сил и объемные заряды, запишутся в виде

$$\text{rot } \mathbf{H} = \sigma \mathbf{E} + \epsilon_r \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}; \quad (11.77)$$

$$\text{rot } \mathbf{E} = -\mu_r \mu_0 \frac{\partial \mathbf{H}}{\partial t}; \quad (11.78)$$

$$\text{div } \mathbf{H} = 0; \quad (11.79)$$

$$\text{div } \mathbf{E} = 0. \quad (11.80)$$

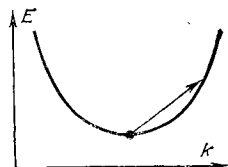


Рис. 11.27. Переход свободного электрона в зоне проводимости

В этих уравнениях σ — удельная проводимость вещества при данной оптической частоте; μ_r , μ_0 — относительная магнитная проницаемость и магнитная постоянная, а ϵ_r , ϵ_0 — относительная диэлектрическая проницаемость и электрическая постоянная соответственно.

Из уравнений (11.77) и (11.78) следует:

$$\begin{aligned} \operatorname{rot} \operatorname{rot} \mathbf{E} &= \operatorname{rot} \left[-\mu_r \mu_0 \frac{\partial \mathbf{H}}{\partial t} \right] = -\mu_r \mu_0 \frac{\partial}{\partial t} \operatorname{rot} \mathbf{H} = \\ &= -\mu_r \mu_0 \left(\sigma \frac{\partial \mathbf{E}}{\partial t} + \epsilon_r \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \right). \end{aligned} \quad (11.81)$$

Учитывая (11.80), имеем:

$$\begin{aligned} \operatorname{rot} \operatorname{rot} \mathbf{E} &= [\nabla (\nabla \mathbf{E})] = \nabla (\nabla \mathbf{E} - \nabla^2 \mathbf{E}) = \\ &= \operatorname{grad} \operatorname{div} \mathbf{E} - \nabla^2 \mathbf{E} = -\nabla^2 \mathbf{E}. \end{aligned} \quad (11.82)$$

Следовательно,

$$\nabla^2 \mathbf{E} - \mu_r \mu_0 \frac{\partial \mathbf{E}}{\partial t} - \mu_r \mu_0 \epsilon_r \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \quad (11.83)$$

Пусть электромагнитная волна, падающая на полупроводник, распространяется вдоль направления x со скоростью v . Для этого случая решением уравнения (11.83) будет

$$\mathbf{E} = \mathbf{E}_0 e^{i\omega(t - x/v)} \quad (11.84)$$

при условии, если

$$1/v^2 = \mu_r \mu_0 \epsilon_r \epsilon_0 - \frac{i\mu_r \mu_0 \sigma}{\omega}. \quad (11.85)$$

Но скорость распространения света в полупроводнике с комплексным показателем преломления

$$\hat{n} = \bar{n} - i\chi \quad (11.86)$$

связана со скоростью распространения света в вакууме формулой

$$v = c/\hat{n}. \quad (11.87)$$

Здесь χ — мнимая часть показателя преломления, которую называют коэффициентом экстинкции или показателем поглощения.

Следовательно,

$$\hat{n}^2 = c^2 \mu_r \mu_0 \left(\epsilon_r \epsilon_0 - \frac{i\sigma}{\omega} \right). \quad (11.88)$$

Если учесть, что для вакуума $\hat{n} = 1$, $\chi = 0$ и $\epsilon_r = 1$ и $\sigma = 0$, то $c^2 \mu_0 \epsilon_0 = 1$ или $c^2 = (\mu_0 \epsilon_0)^{-1}$. Кроме того, при определении оптических свойств полупроводников можно положить $\mu_r = 1$. В соответствии с этим (11.88) запишется в виде

$$\hat{n}^2 = \epsilon - \frac{i\sigma}{\epsilon_0 \omega}. \quad (11.89)$$

На основе (11.86) и (11.89) имеем:

$$\hat{n}^2 = \bar{n}^2 - \chi^2 - i2\bar{n}\chi = \epsilon_r - i\sigma/\epsilon_0 \omega, \quad (11.90)$$

откуда следует, что

$$\bar{n}^2 = \chi^2 + \varepsilon_r; \quad (11.91)$$

$$\chi = \sigma/2\bar{n}\varepsilon_0\omega. \quad (11.92)$$

Поскольку на основе равенств (11.87) и (11.86)

$$\frac{1}{v} = \frac{\bar{n}}{c} - \frac{i\chi}{c}, \quad (11.93)$$

то, подставляя это выражение в (11.84), получаем:

$$\mathcal{E} = \mathcal{E}_0 e^{i\omega(t - \bar{n}x/c)} = \mathcal{E}_0 e^{-\omega\chi x/c} e^{i\omega(t - \bar{n}x/c)}. \quad (11.94)$$

В этой формуле множитель $e^{-\omega\chi x/c}$ выражает собой затухание волны. Так как падающая мощность излучения равна

$$P(0) = J\mathcal{E}(0) = \sigma\mathcal{E}^2(0) = \sigma\mathcal{E}_0^2 e^{i2\omega(t - \bar{n}x/c)},$$

а мощность излучения, оставшаяся после прохождения в полупроводнике расстояния x , есть

$$P(x) = \sigma\mathcal{E}^2(x) = \sigma\mathcal{E}_0^2 e^{-2\omega\chi x/c} e^{i2\omega(t - \bar{n}x/c)}, \quad (11.94a)$$

то прошедшая мощность излучения по отношению к падающей составляет

$$\frac{P(x)}{P(0)} = e^{-2\omega\chi x/c}. \quad (11.95)$$

Но согласно (11.4) эту величину можно выразить через коэффициент поглощения

$$\frac{P(x)}{P(0)} = e^{-\alpha x}. \quad (11.96)$$

Из сопоставления (11.95) и (11.96), а также с учетом (11.92) получим:

$$\alpha = \frac{2\omega\chi}{c} = \frac{\sigma}{c\bar{n}\varepsilon_0}. \quad (11.97)$$

Таким образом, коэффициент поглощения свободными носителями заряда зависит от удельной проводимости вещества.

Поскольку проводимость $\sigma = en\mu_n = \frac{e^2 \langle \tau \rangle}{m_n^*} n$, то коэффициент поглощения запишется через концентрацию свободных носителей заряда в виде

$$\alpha = \frac{\sigma}{c\bar{n}\varepsilon_0} = \frac{e^2 \langle \tau \rangle}{c\bar{n}\varepsilon_0 m_n^*} n. \quad (11.98)$$

Следовательно, коэффициент поглощения свободными носителями заряда зависит от их концентрации n и эффективной массы m^* , среднего времени релаксации $\langle \tau \rangle$ и показателя преломления среды $\bar{n}$ (действительная часть показателя). Но показатель преломления зависит от длины волны, поэтому и коэффициент поглощения свободными носителями заряда должен зависеть от длины волны.

Для того чтобы определить зависимость удельной проводимости σ или показателя поглощения χ от длины падающего света λ , рассмотрим поляризацию вещества. Рассмотрение проведем для случая, когда длина волны падающего излучения находится в области длин волн, намного превышающих длину волны края собственного поглощения. Под действием поля $\mathcal{E}e^{i\omega t}$ такой падающей электромагнитной волны будут смещаться из положения равновесия только свободные электроны. Одновременно на них действует тормозящая сила, пропорциональная скорости, т. е. dx/dt . Поэтому уравнение движения электрона проводимости в электрическом поле падающей волны запишется как

$$m_n^* \frac{d^2x}{dt^2} + m_n^* g \frac{dx}{dt} = e\mathcal{E}e^{i\omega t}. \quad (11.99)$$

Решение этого уравнения будем искать в виде

$$x = Be^{i\omega t}. \quad (11.100)$$

Подставляя (11.100) в (11.99), будем иметь:

$$-m_n^* \omega^2 B + im_n^* g B \omega = e\mathcal{E}, \quad (11.101)$$

откуда находим:

$$B = \frac{e\mathcal{E}}{m_n^* \omega (-\omega + ig)}. \quad (11.102)$$

Если в единице объема вещества имеется n свободных электронов, то поляризуемость ее составляет $P = enB/\mathcal{E}$, а относительная диэлектрическая проницаемость равна $\epsilon_r = \epsilon_0 + P$. Таким образом,

$$\epsilon_r = \epsilon_0 + \frac{e^2 n}{m_n^* \omega (-\omega + ig)}. \quad (11.103)$$

Но комплексный показатель преломления среды есть $\tilde{n} = (\epsilon_r/\epsilon_0)^{1/2}$, поэтому

$$(\tilde{n} - i\chi)^2 = \frac{\epsilon_r}{\epsilon_0} = 1 + \frac{e^2 n}{\epsilon_0 m_n^* \omega (-\omega + ig)}, \quad (11.104)$$

откуда следует, что

$$\tilde{n}^2 - \chi^2 - 1 = -\frac{e^2 n}{\epsilon_0 m_n^* (\omega^2 + g^2)}; \quad (11.105)$$

$$2\tilde{n}\chi = \frac{e^2 ng}{\epsilon_0 m_n^* \omega (\omega^2 + g^2)}. \quad (11.106)$$

В случае стационарного поля, для которого из (11.99)

$$m_n^* g \frac{dx}{dt} = e\mathcal{E}, \quad (11.107)$$

а скорость движения электрона проводимости

$$v = \frac{dx}{dt} = \mu_n \mathcal{E}, \quad (11.108)$$

имеем

$$g = \frac{e\mathcal{E}}{m_n^*v} = \frac{e}{m_n^*\mu_n}. \quad (11.109)$$

Поэтому соотношение (11.106) можно записать в виде

$$2\bar{n}\chi = \frac{en\mu_n}{\varepsilon_0\omega \left[1 + \left(\frac{m_n^*\mu_n\omega}{e} \right)^2 \right]} = \frac{\sigma}{\varepsilon_0\omega \left[1 + \left(\frac{m_n^*\mu_n\omega}{e} \right)^2 \right]}. \quad (11.110)$$

Соответственно

$$\alpha = \frac{2\omega\chi}{c} = \frac{\sigma}{c\bar{n}\varepsilon_0 \left[1 + \left(\frac{m_n^*\mu_n\omega}{e} \right)^2 \right]}. \quad (11.111)$$

Если $\frac{m_n^*\mu_n\omega}{e} \gg 1$ и учитывая, что $\omega = 2\pi\nu = \frac{2\pi c}{\lambda}$, из (11.111) следует:

$$\alpha = \frac{e^3 n}{4\pi^2 c^3 \varepsilon_0 \bar{n} m_n^{*2} \mu_n} \lambda^2 = \frac{e^2 n}{4\pi^2 c^3 \varepsilon_0 \bar{n} m_n^* \langle \tau \rangle} \lambda^2. \quad (11.112)$$

Как видно, коэффициент поглощения свободными носителями заряда пропорционален их концентрации, квадрату длины волны падающего света и обратно пропорционален эффективной массе и времени релаксации.

Зависимость α от λ^2 хорошо выполняется для сурьмянистого индия n -типа, которая представлена на рис. 11.28. Кривая на рис. 11.28, а соответствует концентрации электронов $3,5 \cdot 10^{17} \text{ см}^{-3}$, кривая на рис. 11.28, б — концентрации $6,2 \cdot 10^{17} \text{ см}^{-3}$.

Поскольку время релаксации $\langle \tau \rangle$ зависит от природы рассеивающих центров, то можно ожидать, что на зависимость коэффициента поглощения свободными носителями заряда от длины волны падающего света будет оказывать влияние механизм рассеяния. Действительно, рассеяние на акустических фонах приводит к поглощению, меняющемуся как $\lambda^{1,5}$, рассеяние на оптических фонах определяет

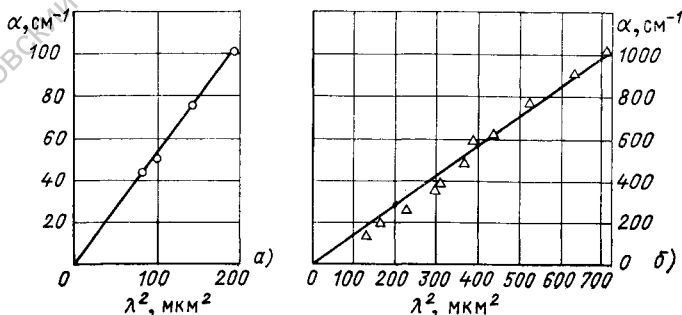


Рис. 11.28. Зависимость коэффициента поглощения InSb n -типа от квадрата длины волны для концентрации электронов $3,5 \cdot 10^{17}$ (а) и $6,2 \cdot 10^{17} \text{ см}^{-3}$ (б)

зависимость $\lambda^{2,5}$, а рассеяние ионизованными примесями дает зависимость $\lambda^{3,5}$. В общем случае, когда реализуются все типы рассеяния, коэффициент поглощения свободными носителями заряда равен сумме трех членов

$$\alpha = A\lambda^{1,5} + B\lambda^{2,5} + C\lambda^{3,5}, \quad (11.113)$$

где A, B, C — константы.

В зависимости от концентрации легирующих примесей, а следовательно, от концентрации свободных электронов будет доминирующим тот или иной механизм рассеяния и $\alpha \sim \lambda^p$. В табл. 11.3 приведены значения p для различных полупроводников.

Т а б л и ц а 11.3

Полупроводник	Концентрация электронов, 10^{17} см^{-3}	p	Полупроводник	Концентрация электронов, 10^{17} см^{-3}	p
Ge	0,5—5	2	InP	0,4—4	2,5
GaP	10	1,8	InAs	0,3—8	3
GaAs	1—5	3	InSb	1—3	2
GaSb	0,5	3,5	AlSb	0,4—4	2

У большинства полупроводников валентная зона состоит из трех подзон, разделенных за счет спин-орбитального взаимодействия (такая зонная структура показана на рис. 11.29). Поэтому в полупроводниках p -типа, где вершина валентной зоны занята дырками, возможно оптическое поглощение, обусловленное переходами электронов между отдельными подзонами: a — из зоны легких дырок V_2 в зону тяжелых дырок V_1 ; b — из отщепленной зоны V_3 в зону тяжелых дырок V_1 ; c — из отщепленной зоны V_3 в зону легких дырок V_2 . Для таких прямых переходов соблюдаются правила отбора.

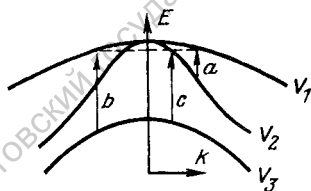


Рис. 11.29. Структура валентных подзон и внутренние переходы

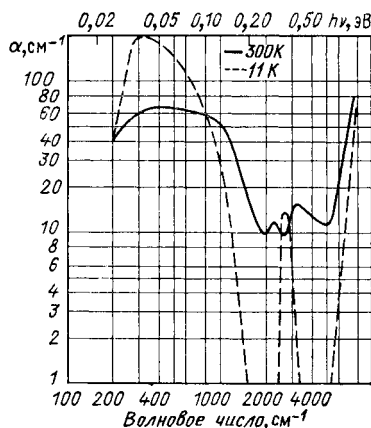


Рис. 11.30. Поглощение, связанное с переходами между валентными подзонами в германии p -типа

Интенсивность и положение полос поглощения должны меняться с изменением степени легирования полупроводника или температуры. Если ширина запрещенной зоны полупроводника меньше энергии спин-орбитального расщепления, то переходы $V_3 \rightarrow V_1$ и $V_3 \rightarrow V_2$ будут трудно различимы на фоне собственного поглощения. В качестве иллюстрации на рис. 11.30 приведено поглощение, связанное с переходами электронов между валентными подзонами в германии p -типа.

11.8. ПРИМЕСНОЕ ПОГЛОЩЕНИЕ

Поглощение света, приводящее к ионизации или возбуждению примесных центров в кристалле, называют примесным поглощением. Допустим, что в кристаллической решетке полупроводника имеется донорная или акцепторная примесь (рис. 11.31, а, б) или примесь с глубокими энергетическими уровнями (рис. 11.31, в, г). У такого полупроводника будет наблюдаться полоса поглощения, вызываемая переходами электронов с энергетических уровней примеси в зону проводимости или переходами электронов валентной зоны на примесные уровни. Так как энергия ионизации примесного центра меньше энергии, требуемой для перевода электрона из валентной зоны в зону проводимости, полосы примесного поглощения располагаются за краем фундаментального поглощения.

Кроме основных состояний примесные центры могут иметь и энергетические уровни возбуждения, энергия которых для доноров и акцепторов согласно (2.174) определяется выражением

$$E = 13,52 \frac{Z^2}{\epsilon_r^2} \left(\frac{m^*}{m_0} \right) \frac{1}{n^2}, \quad (11.114)$$

где $n = 1, 2, \dots$

Оптическое возбуждение электронов примеси, приводящее к переводу электрона примеси из основного состояния в возбужденное, будет сопровождаться поглощением, в спектре которого будет наблюдаться несколько полос. В качестве примера на рис. 11.32 представлен спектр поглощения кремния, легированного бором с $N_a = 1,8 \cdot 10^{16} \text{ см}^{-3}$, при 4,2 К.

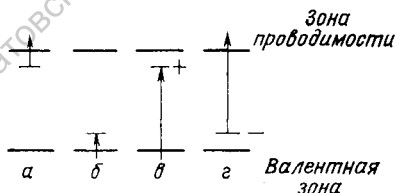


Рис. 11.31. Поглощение, связанное с переходами между примесными уровнями и зонами:

а — донор — зона проводимости; б — валентная зона — акцептор; в — валентная зона — донор; г — акцептор — зона проводимости

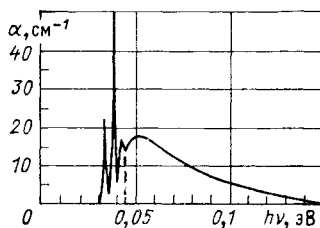


Рис. 11.32. Спектр поглощения кремния, легированного бором

11.9. РЕШЕТОЧНОЕ ПОГЛОЩЕНИЕ

Бинарные полупроводники $A^{III}B^V$ состоят из атомов различного типа и имеют ионную составляющую химической связи. Поэтому их можно рассматривать как состоящие из электрических диполей. Эти диполи могут взаимодействовать с электромагнитным излучением. Наиболее сильное поглощение электромагнитного излучения имеет место, когда частота излучения равна частоте собственных колебаний диполей. Поглощение света полупроводником, связанное с возбуждением колебаний кристаллической решетки, называют *решеточным поглощением*. Так как решетка кристалла может поглощать энергию электромагнитного поля излучения только при определенных значениях энергии фотона, спектр ее характеризуется рядом пиков поглощения, лежащих в далекой инфракрасной области спектра (рис. 11.33).

Значения частот излучения, поглощаемого колебаниями решетки, определяются совокупностью оптических и акустических ветвей колебаний. В полупроводниках имеются две поперечные оптические (ТО) ветви колебаний, две поперечные акустические (ТА), одна продольная оптическая (ЛО) и одна продольная акустическая (ЛА).

При поглощении света решеткой должен выполняться закон сохранения квазиимпульса, т. е. должны поглощаться только такие фотоны, импульс которых равен квазиимпульсу фонона. Но импульс фонона $q = h/\lambda_{\text{фон}}$ мал по сравнению с квазиимпульсом фотона $P = h/\lambda_{\text{фот}}$, поэтому для выполнения закона сохранения квазиимпульса должно быть испущено два или несколько фононов. Все это обуславливает сложную структуру спектра поглощения света колебаниями решетки. На рис. 11.33 приведена часть спектра решеточного поглощения электронного арсенида галлия, а идентификация этих пиков поглощения дана в табл. 11.4.

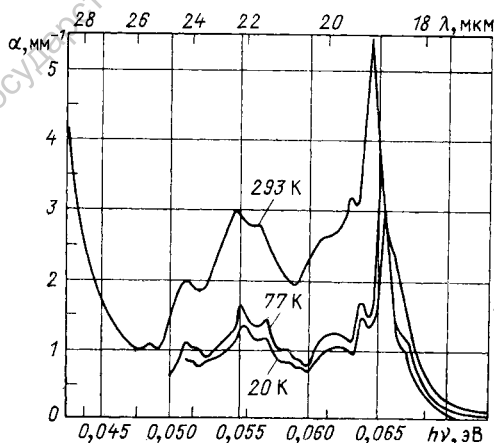


Рис. 11.33. Спектр решеточного поглощения GaAs в области энергии 0,04—0,07 эВ при разных температурах

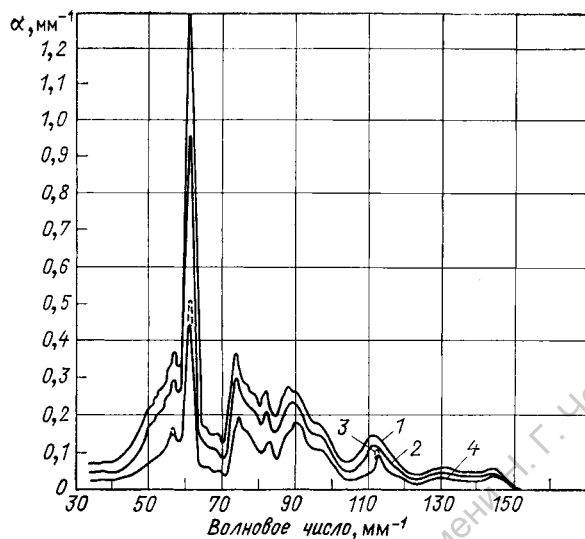


Рис. 11.34. Решеточное поглощение кремния при разных температурах:
1 — 365 К; 2 — 290 К; 3 — 77 К; 4 — 20 К

Таблица 11.4. Фононы, участвующие в решеточном поглощении арсенида галлия

Положение максимума, эВ	Тип и энергия фононов, эВ	Положение максимума, эВ	Тип и энергия фононов, эВ
0,0648	$\left\{ \begin{array}{l} TO_1 + TO_1 \\ 0,0324 + 0,0324 \end{array} \right.$	0,058	$\left\{ \begin{array}{l} LO + LO \\ 0,029 + 0,029 \end{array} \right.$
0,0631	$\left\{ \begin{array}{l} TO_2 + TO_2 \\ 0,0316 + 0,0316 \end{array} \right.$	0,0565	$\left\{ \begin{array}{l} TO_1 + LA \\ 0,0324 + 0,0241 \end{array} \right.$
	$\left\{ \begin{array}{l} TO_1 + LO \\ 0,0324 + 0,0288 \end{array} \right.$	0,0548	$\left\{ \begin{array}{l} TO_2 + LA \\ 0,0316 + 0,0232 \end{array} \right.$
0,0612	$\left\{ \begin{array}{l} TO_2 + LO \\ 0,0316 + 0,0296 \end{array} \right.$	0,0510	$\left\{ \begin{array}{l} LO + LA \\ 0,0288 + 0,0222 \end{array} \right.$

Таблица 11.5. Фононы, участвующие в решеточном поглощении кремния

Волновое число, мм⁻¹	Энергия пика, эВ	Тип фононов
144,8	0,1795	3TO
137,8	0,1708	2TO + LO
130,2	0,1614	2TO + LO
96,4	0,1195	2TO
89,6	0,1111	TO + LO
81,9	0,1015	TO + LA
74,0	0,0917	LO + LA
68,9	0,0756	TO + TA
61,0	0,0702	LO + TA

Здесь $TO = 0,0598$ эВ; $LO = 0,0513$ эВ; $LA = 0,0414$ эВ; $TA = 0,0158$ эВ.

В гомополярных полупроводниках, таких как кремний и германий, нет диполей. Но их атомы под воздействием поля световой волны приобретают дипольный момент. Индуцируемый диполь сильно взаимодействует с излучением и в результате возникают фононы. Решеточное поглощение германия лежит в инфракрасной области спектра от 14 до 33 мкм, а кремния — в области от 7 до 30 мкм. На рис. 11.34 показан спектр решеточного поглощения кремния, а идентификация фононов приведена в табл. 11.5.

ГЛАВА ДВЕНАДЦАТАЯ

ЛЮМИНЕСЦЕНЦИЯ ПОЛУПРОВОДНИКОВ

12.1. ТИПЫ ЛЮМИНЕСЦЕНЦИИ

У полупроводников, находящихся в возбужденном состоянии, может наблюдаться испускание электромагнитного излучения — люминесценция, которая является процессом, обратным поглощению. Такое состояние вещества можно получить различными способами. При фотолюминесценции излучение появляется за счет поглощаемой энергии света. Свечение, возникающее при бомбардировке образца электронами, называют катодолюминесценцией. Люминесценция может наблюдаться при возбуждении вещества рентгеновским излучением (рентгенолюминесценция) под действием электрического поля (электролюминесценция) или в результате механического воздействия на вещество (триболюминесценция), а также при химических и биологических процессах (хеми- и биолюминесценция).

Согласно определению, впервые сформулированному С. И. Вавиловым, *люминесценция есть свечение, избыточное над тепловым излучением тела, если это избыточное излучение обладает конечной длительностью, значительно превышающей период световых колебаний*. Данное определение отделяет люминесценцию от теплового равновесного излучения тела и дает право отнести ее к неравновесным излучениям.

Люминесценция, как любое электромагнитное излучение, характеризуется интенсивностью, спектральным составом, поляризацией, когерентностью и длительностью послесвечения. И только длительность послесвечения отличает люминесценцию от всех других видов неравновесного излучения, таких как отражение и рассеяние света, тормозное излучение, излучение Вавилова — Черенкова, которые практически безынерционны; их длительность составляет примерно период световых колебаний (10^{-15} с). В случае люминесценции акты поглощения и излучения света разделены промежуточными процессами, что и приводит к продолжительному существованию свечения после прекращения возбуждения.

У твердых тел различают в основном два вида люминесценции: мономолекулярную и рекомбинационную.

12.2. МОНОМОЛЕКУЛЯРНОЕ СВЕЧЕНИЕ ТВЕРДЫХ ТЕЛ

Мономолекулярное излучение твердых тел имеет место в том случае, когда поглощение и излучение света происходят внутри примесного центра.

Спектры люминесценции твердых тел, т. е. зависимость интенсивности свечения от энергии излученного фотона или длины волны, отличаются от соответствующих спектров изолированных атомов примеси двумя основными чертами. Во-первых, излучение твердых тел наблюдается обычно в виде широких полос и, во-вторых, их спектр люминесценции, как правило, смещен в сторону направлений длинных волн по сравнению со спектром поглощения, т. е. излучаемая энергия меньше, чем поглощаемая. Разницу

этих двух энергий называют «сдвигом Франка — Кондона». Эти особенности оптических свойств твердых тел обязаны своим происхождением взаимодействию центра свечения с полем кристаллической решетки. При взаимодействии центра с ближайшим окружением его потенциальная энергия складывается из кулоновского притяжения, обменного взаимодействия и ван-дер-ваальсовского притяжения. В результате этих сил взаимодействия атомные уровни примеси основного U_0 и возбужденного U_1 состояний в кристаллической решетке будут квазимолекулярными, как показано на рис. 12.1.

При такой ситуации возможно поглощение света примесным центром, в результате которого электрон с основного уровня U_0 перейдет на возбужденный U_1 . Самопроизвольный обратный переход электрона с возбужденного уровня центра на основной будет сопровождаться излучением. Если учесть заселенность колебательных уровней электронами в основном и возбужденном состояниях, то совокупность соответствующих фотонов образует полосы поглощения и излучения, которые на рис. 12.1 изображены на оси ординат — шкале энергии. Длительность такой люминесценции твердого тела будет определяться временем жизни электрона в возбужденном состоянии центра. Оно постоянно, не зависит от внешних факторов и, в частности, для дипольного излучения атомов составляет примерно 10^{-8} с.

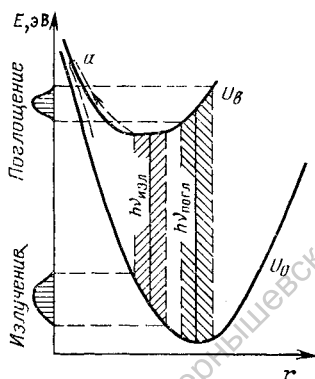


Рис. 12.1. Схема энергетических уровней примесного центра в твердом теле

12.3. РЕКОМБИНАЦИОННОЕ ИЗЛУЧЕНИЕ ПОЛУПРОВОДНИКОВ ПРИ ФУНДАМЕНТАЛЬНЫХ ПЕРЕХОДАХ

При возбуждении полупроводника, создающем электронно-дырочные пары, рекомбинационное излучение может возникнуть при непосредственной рекомбинации свободных электронов и свободных дырок, при аннигиляции экситона, а также при рекомбинации сво-

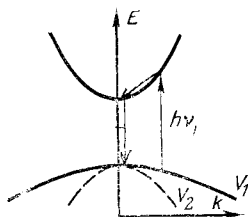


Рис. 12.2. Прямые излучательные переходы

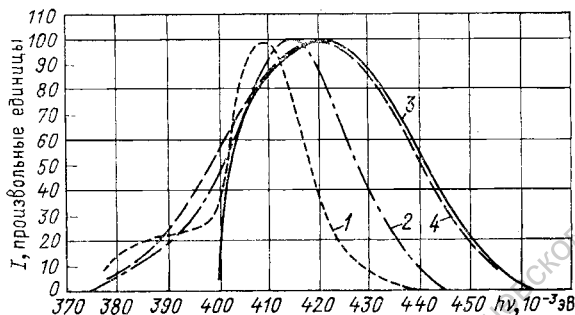


Рис. 12.3. Излучение антимоноида индия n -типа при 77 К 1 — $2,3 \cdot 10^{16} \text{ см}^{-3}$; 2 — $9 \cdot 10^{16} \text{ см}^{-3}$; 3 — $1,8 \cdot 10^{17} \text{ см}^{-3}$ (теория); 4 — $1,8 \cdot 10^{17} \text{ см}^{-3}$

бодных электронов и свободных дырок через рекомбинационные ловушки, которые выступают здесь в качестве центров свечения. Сначала рассмотрим фундаментальные переходы, при которых изменение энергии равно ширине запрещенной зоны или близко к ней.

1. Прямые переходы «зона проводимости — валентная зона»

В полупроводнике с прямой запрещенной зоной поглощение света сопровождается образованием свободного электрона и свободной дырки, волновые векторы которых одинаковы ($k = k'$). После процесса поглощения света образовавшиеся свободные носители заряда претерпевают рассеяние, в результате чего за время релаксации электрон опускается на дно зоны проводимости, а дырка поднимается к потолку валентной зоны. Такое состояние между свободными носителями заряда и кристаллической решеткой устанавливается за время около 10^{-10} — 10^{-12} с. Прямая рекомбинация свободного электрона и свободной дырки с излучением фотона наиболее вероятна, если после процесса релаксации волновые векторы электрона и дырки будут одинаковы (рис. 12.2). В этом случае, как и при соответствующих процессах поглощения [формула (11.24)], спектр излучения описывается выражением

$$I(h\nu) = B(h\nu - E_g)^{1/2}, \quad (12.1)$$

где B — постоянная.

Из этой формулы следует, что излучение при рекомбинации свободных носителей заряда должно иметь порог со стороны низких энергий при $h\nu = E_g$ (рис. 12.3, кривая 3). При увеличении температуры, а также при возрастании энергии возбуждения заполняются более высокие состояния в зоне проводимости. Это обусловит излучение при более высоких энергиях фотонов, приводящее к появлению в спектре люминесценции высокоэнергетического «хвоста».

В качестве примера прямой излучательной межзонной рекомбинации на рис. 12.3 представлен спектр фотолюминесценции InAs

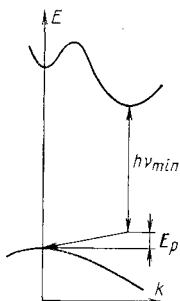


Рис. 12.4. Непрямые излучательные переходы

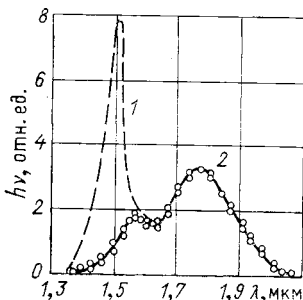


Рис. 12.5. Спектр излучения германия (толщина образца $1,3 \times 10^{-3}$ см):

кривая 2 — экспериментальная;
кривая 1 получена с учетом реабсорбции (самопоглощения).

n -типа с различным уровнем легирования. Сдвиг максимума излучения и его высокоэнергетического края в сторону больших энергий фотонов при возрастании легирования обусловлен смещением уровня Ферми в зону проводимости.

2. Непрямые переходы «зона проводимости — валентная зона»

В полупроводнике с непрямой запрещенной зоной поглощение света происходит с поглощением или излучением фонона. В таком полупроводнике при рекомбинации свободного электрона и свободной дырки должны также принимать участие фононы, которые обеспечивают сохранение квазиимпульса (рис. 12.4). Наиболее вероятным процессом является эмиссия фонона. Процесс поглощения фонона по сравнению с излучением фонона несуществен, поскольку его вероятность в $N_p/(N_p + 1)$ раз меньше вероятности испускания фонона. Оптический переход, сопровождаемый испусканием фонона, имеет меньшую энергию, чем ширина запрещенной зоны, равную $h\nu_{min} = E_g - E_p$. В таком случае по аналогии с коэффициентом поглощения [формула (11.35)] спектр излучения описывается выражением

$$I(h\nu) = B(h\nu - E_g + E_p)^2, \quad (12.2)$$

где E_p — энергия фонона, участвующего в рекомбинации.

Если у полупроводника возможны прямые и не прямые переходы, в спектре его излучения наблюдаются две полосы люминесценции, как, например, в германии (рис. 12.5). Длинноволновая полоса излучения определяется не прямой излучательной рекомбинацией. Коротковолновая полоса люминесценции обусловлена переходами электронов с излучением из прямой долины, которая расположена на 0,15 эВ выше дна самой низкой долины. Кривая 2 на этом рисунке — экспериментальная, а кривая 1 получена из кривой 2 путем введения поправки на поглощение в образце.

3. Сильно легированный полупроводник

У сильно легированного полупроводника уровень Ферми лежит внутри разрешенной зоны (в зоне проводимости для n -типа и в валентной зоне для p -типа). На рис. 12.6 приведена энергетическая диаграмма полупроводника, в котором за счет возбуждения создана концентрация электронов в зоне проводимости.

Для полупроводника с прямой запрещенной зоной для прямых переходов интенсивность излучения с увеличением частоты должна расти как

$$I(h\nu) = B(h\nu - E_g)^{1/2}. \quad (12.3)$$

В случае непрямых переходов, при которых закон сохранения квазимпульса осуществляется за счет процессов электронно-электронного рассеяния или рассеяния «электрон — примесь», все занятые состояния в зоне проводимости могут давать начало излучательным переходам во все пустые состояния валентной зоны. В этом случае одной и той же энергии $h\nu$ могут отвечать переходы между различными верхними и нижними состояниями, поэтому спектр излучения описывается выражением

$$I(h\nu) = B(h\nu - E_g)^2. \quad (12.4)$$

4. Экситонная рекомбинация

В достаточно чистом полупроводнике при возбуждении светом, энергия которого близка к ширине запрещенной зоны, возможно возникновение экситона — свободного электрона и свободной дырки, связанных силой кулоновского притяжения. При рекомбинации таких носителей заряда, т. е. при аннигиляции экситона, будет излучаться узкая спектральная линия. В полупроводнике с прямой запрещенной зоной, имеющем разрешенные переходы (рис. 12.7, а), энергия экситонного излучения равна:

$$h\nu = E_g - E_{\text{экс}}, \quad (12.5)$$

где $E_{\text{экс}}$ — энергия связи экситона.

Экситон может иметь возбужденные состояния, энергия ионизации которых согласно (11.71) в n^2 раз меньше, чем энергия ионизации основного состояния, соответствующего $n = 1$. Поэтому излучение свободного экситона может состоять из серии узких линий.

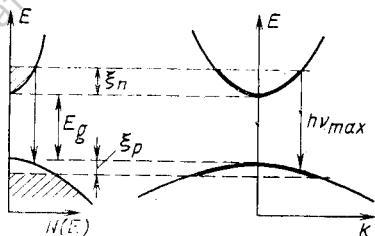


Рис. 12.6. Прямые переходы в случае возбуждения

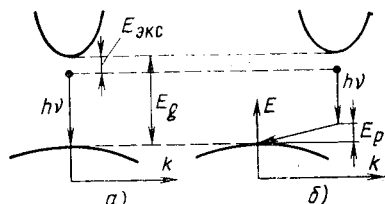


Рис. 12.7. Экситонная рекомбинация а — прямая; б — непрямая

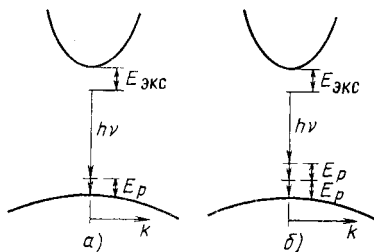
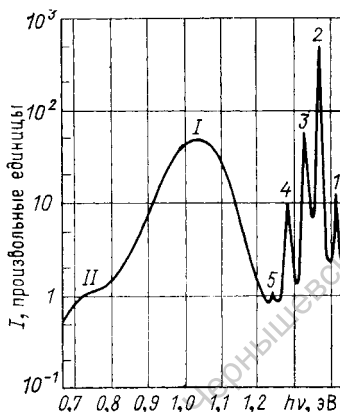


Рис. 12.8. Прямая экситонная рекомбинация:
а — с испусканием одного оптического фонона; б — с испусканием двух оптических фононов

Рис. 12.9. Спектр фотолюминесценции InP при 6 К



В полупроводнике с непрямой запрещенной зоной (рис. 12.7, б), у которого закон сохранения квазиимпульса осуществляется за счет испускания оптического фонона с энергией E_p , энергия излученного фотона

$$h\nu = E_g - E_{\text{экс}} - E_p. \quad (12.6)$$

У полупроводников с прямыми разрешенными переходами излучение свободного экситона может также происходить с испусканием одного или нескольких фононов (рис. 12.8), но вероятность таких переходов уменьшается и соответствующая линия излучения становится слабее.

В присутствии примесей могут образоваться связанные экситоны. Их спектр излучения состоит из узких линий с энергиями, меньшими, чем при рекомбинации свободного экситона. На рис. 12.9 в качестве примера приведен спектр фотолюминесценции достаточно чистого фосфида индия. Пик излучения, обозначенный цифрой 1, относится к рекомбинации свободного экситона, а линии 2, 3, 4 и 5 обусловлены рекомбинацией связанного экситона с испусканием соответственно 0, 1, 2 и 3 продольных оптических фононов LO , энергия которого составляет 43 мэВ. Полосы излучения I и II вызваны процессами неэкситонного характера.

12.4. РЕКОМБИНАЦИОННОЕ ИЗЛУЧЕНИЕ ПРИ ПЕРЕХОДАХ МЕЖДУ ЗОНОЙ И ПРИМЕСНЫМИ УРОВНЯМИ

В примесных полупроводниках, в которых в качестве примеси могут выступать не только доноры и акцепторы, но и разного рода дефекты, нарушения кристаллической периодичности на поверхности полупроводника, сверхстехиометрический избыток одного из компонентов в случае сложного полупроводникового материала, важную роль в процессах люминесценции играют центры, энергетические уровни которых попадают в запрещенную зону полупроводника. В таких случаях люминесценция идет преимущественно через центры

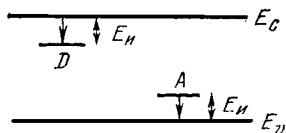


Рис. 12.10. Мелкие переходы

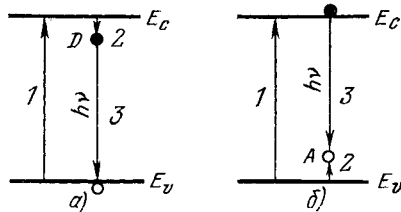


Рис. 12.11. Излучательные переходы между зоной и примесными состояниями

рекомбинации. Центры электронно-дырочной рекомбинации характеризуются эффективными сечениями захвата электронов σ_n и дырок σ_p . Ловушки и донорные примеси имеют большую вероятность захвата электронов из зоны проводимости, однако переход электронов с этих центров в валентную зону практически невозможен, т. к. $\sigma_n \gg \sigma_p$. Если центр не имеет электрона, то это — электронная ловушка, а если центр имеет электрон, то центр является донором.

Переходы при нейтрализации ионизованных доноров и акцепторов (рис. 12.10) в принципе могут являться излучательными и дают свечение в далекой инфракрасной области спектра. Однако, как показали расчеты, вероятность переходов с испусканием фононов значительно больше, чем вероятность излучательных переходов. Например, для германия n -типа при 4 К сечение захвата для переходов с испусканием фонона равно примерно 10^{-12} см², а сечение захвата при переходе с излучением фотона составляет $4 \cdot 10^{-19}$ см². Процесс безызлучательной рекомбинации происходит следующим образом. Сначала электрон зоны проводимости захватывается на возбужденный уровень донора. Затем он совершает каскадные переходы на нижележащие энергетические уровни, испуская фонон при каждом переходе.

1. Переход «зона — уровень примеси»

У полупроводников n - и p -типа, как следует из рис. 12.11, возможны излучательные переходы с донорного уровня в валентную зону или из зоны проводимости на акцепторный уровень. Такие центры отличаются от ловушек лишь несколько большей глубиной залегания уровней и тем, что захват дырки из валентной зоны хотя и менее вероятен по сравнению с захватом электрона из зоны проводимости, но все же имеет заметную величину ($\sigma_n \gg \sigma_p$). Согласно рис. 12.11, а после процесса поглощения (переход 1) сначала происходит захват электрона из зоны проводимости центром (переход 2), а затем осуществляется излучательная рекомбинация локализованного электрона и дырки валентной зоны (переход 3).

В другой модели люминесценции (рис. 12.11, б) локализованный уровень расположен немного выше валентной зоны и имеет большую вероятность захвата дырки ($\sigma_p \gg \sigma_n$), но вместе с тем имеет и заметную вероятность захвата электронов зоны проводимости. Люмине-

Рис. 12.12. Спектр катодолюминесценции при 4,2 К арсенида галлия, легированного цинком (см^{-3})

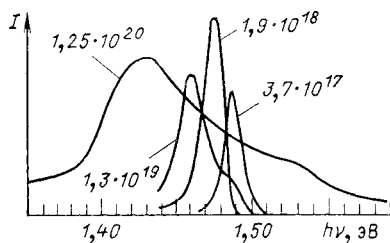
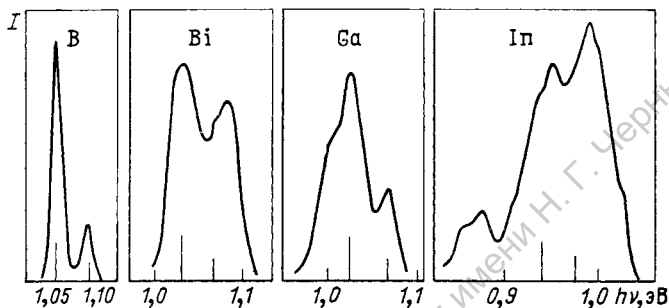


Рис. 12.13. Спектральное распределение примесной излучательной рекомбинации в кремнии при 20 К



сцентное излучение возникает при рекомбинации свободного электрона зоны проводимости с захваченной этим уровнем дыркой (переход 3).

При прямых переходах излучается фотон с энергией

$$h\nu = E_g - E_{\text{и}}, \quad (12.7)$$

а при непрямых переходах энергия излученного фотона равна:

$$h\nu = E_g - E_{\text{и}} - E_p, \quad (12.8)$$

где $E_{\text{и}}$ — энергия ионизации примесного атома; E_p — энергия фона, принимающего участие в рекомбинации.

Такого типа излучательная рекомбинация возможна только для полупроводников с относительно низкой концентрацией примесей. Увеличение концентрации доноров (или акцепторов) вызовет размытие энергетического уровня примеси и превращение его в примесную зону, которая будет смыкаться с краями основной зоны. В качестве примера на рис. 12.12 приведен спектр катодолюминесценции арсенида галлия, легированного цинком. Из рисунка следует, что при возрастании концентрации акцепторов (показано цифрами у кривых) имеет место расширение полосы излучения, что вызывается сужением запрещенной зоны, обусловленным высоким уровнем легирования.

У полупроводников с непрямой запрещенной зоной при малых концентрациях доноров в переходах «донор — валентная зона», а также в переходах «зона проводимости — акцептор» должен принимать участие фотон. На рис. 12.13 приведены спектры излучения для образцов кремния, легированных разными примесями. В ре-

комбинационных переходах принимает участие поперечный акустический фотон TA с энергией $0,016$ эВ и поперечный оптический фотон TO с энергией $0,055$ эВ, которые обеспечивают сохранение квазиимпульса носителя заряда. На рисунке длинная вертикальная черта соответствует энергии $h\nu = E_g - E_n - E_p(TO)$, короткая — энергии $h\nu = E_g - E_n - E_p(TA)$. Значения энергий ионизации для примесей (в электрон-вольтах) составляют: $0,046$ (В); $0,069$ (Ві); $0,071$ (Ga) и $0,16$ (In). У образца, легированного индием, в низкоэнергетической части спектра наблюдаются бесфотонные переходы.

2. Донорно-акцепторные пары

Если в полупроводнике присутствуют донорная и акцепторная примеси с энергией ионизации соответственно E_d и E_a , то они могут образовывать пары, которые будут действовать как неподвижные молекулы, погруженные в кристалл. Кулоновское взаимодействие между донором и акцептором, находящимися в среде с относительной диэлектрической проницаемостью ϵ_r , вызывает уменьшение их энергии связи на величину

$$\Delta E = e^2/\epsilon_r r, \quad (12.9)$$

где r — расстояние между донором и акцептором, образующими пары.

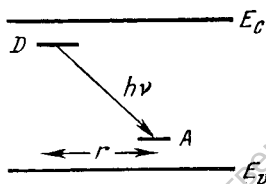
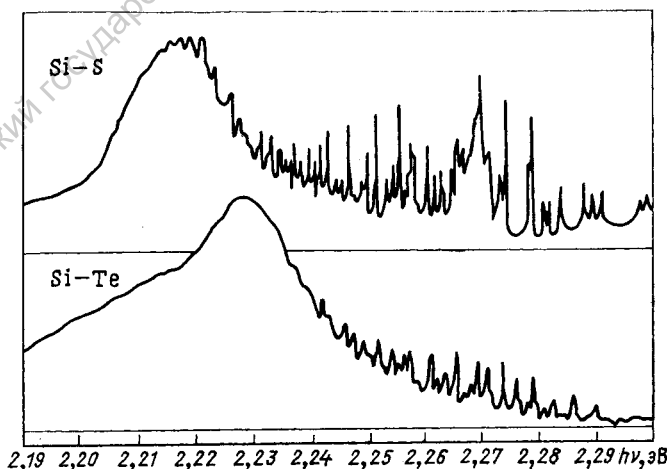


Рис. 12.14. Донорно-акцепторные переходы

Рис. 12.15. Спектры фотолюминесценции кристаллов GaP при $1,6$ К, соответствующие донорно-акцепторной рекомбинации в парах Si—S и Si—Te



Поэтому энергетический зазор между донорным и акцепторным уровнями (рис. 12.14) будет равен:

$$h\nu = E_g - E_d - E_a + \frac{e^2}{\epsilon_r r}. \quad (12.10)$$

Поскольку примеси могут занимать дискретные положения в решетке (например, примесь замещения занимает узлы), расстояние в донорно-акцепторной паре меняется конечными порциями. Это должно обусловить наличие в спектре тонкой структуры. При больших расстояниях между донором и акцептором ($r > 4,0$ нм) линии излучения перекрываются, образуя широкий спектр. Спектральное положение как широкой полосы, так и структуры узких линий должно зависеть от типа примесей. При этом в зависимости от мест, занимаемых примесями, могут образовываться различные типы пар. Если донор и акцептор занимают в решетке одинаковые места, то образуются пары типа I. Если донор и акцептор занимают различные места, то возникают пары типа II.

Излучение донорно-акцепторных пар было обнаружено в ряде полупроводников. В частности, на рис. 12.15 дан спектр фосфида галлия при 1,6 К, имеющего донорно-акцепторные пары Si — S и Si — Te, в которых Si является акцептором. И доноры, и акцепторы занимают узлы вместо фосфора.

12.5. РЕЛАКСАЦИЯ ЛЮМИНЕСЦЕНЦИИ ПОЛУПРОВОДНИКОВ

При выключении источника возбуждения излучение полупроводника исчезает не мгновенно. Поскольку интенсивность люминесценции пропорциональна числу актов рекомбинации в секунду

$$I_{\text{л}} \sim - \left(\frac{dn}{dt} \right)_r, \quad (12.11)$$

то изменение ее яркости во времени после снятия возбуждения будет различным в зависимости от вида люминесценции.

В случае мономолекулярного свечения, когда процесс поглощения и излучения происходит в одном и том же центре, уменьшение концентрации возбужденных электронов во времени определяется соотношением вида

$$n = n_0 e^{-t/\tau}, \quad (12.12)$$

где n_0 — концентрация возбужденных электронов в момент $t = 0$; τ — время жизни электрона на возбужденном уровне центра.

Следовательно, *интенсивность мономолекулярной люминесценции после прекращения возбуждения*

$$I_{\text{л}} \sim e^{-t/\tau}, \quad (12.13)$$

т. е. убывает по экспоненциальному закону с постоянной времени τ .

Для рекомбинационной люминесценции закон затухания определяется степенью ее возбуждения. Рассмотрим два крайних случая. При слабом возбуждении, когда концентрация избыточных носителей

заряда Δn значительно меньше равновесных $\Delta n \ll (n_0 + p_0)$, интенсивность рекомбинации согласно (7.18) пропорциональна первой степени концентрации, т. е.

$$\frac{d\Delta n}{dt} = -\frac{\Delta n}{\tau}. \quad (12.14)$$

Для этого случая $\Delta n = \Delta n_0 e^{-t/\tau}$, поэтому

$$I_{\pi} = I_0 e^{-t/\tau}, \quad (12.15)$$

где Δn_0 и I_0 — концентрация избыточных носителей заряда и интенсивность люминесценции в момент времени $t = 0$, т. е. в момент выключения возбуждающего света.

Вид излучения полупроводника, интенсивность которого описывается (12.15), носит название мономолекулярного рекомбинационного свечения. По его экспоненциальной кривой релаксации можно определить время жизни τ неравновесных носителей заряда.

В случае сильного возбуждения, когда $\Delta n \gg (n_0 + p_0)$, интенсивность рекомбинации пропорциональна квадрату концентрации избыточных носителей заряда [равенство (7.20)], т. е.

$$\frac{d\Delta n}{dt} = -\gamma_r \Delta n^2. \quad (12.16)$$

Поскольку концентрация избыточных носителей убывает во времени как

$$\Delta n = \frac{\gamma_r \Delta n_0}{1 + \gamma_r \Delta n_0 t}, \quad (12.17)$$

то интенсивность люминесценции описывается соотношением вида

$$I_{\pi} = \frac{\gamma_r^2 \Delta n_0^2}{(1 + \gamma_r \Delta n_0 t)^2}, \quad (12.18)$$

а люминесценцию называют бимолекулярным рекомбинационным свечением.

12.6. ТЕМПЕРАТУРНОЕ ТУШЕНИЕ ЛЮМИНЕСЦЕНЦИИ ПОЛУПРОВОДНИКОВ

У полупроводников интенсивность излучения уменьшается при повышении температуры, т. е. наблюдается температурное тушение люминесценции. Если через P_r обозначить вероятность излучательных переходов, которая предполагается не зависящей от температуры, а через P_{nr} — вероятность безызлучательных переходов, то квантовый выход люминесценции η описывается выражением

$$\eta = \frac{P_r}{P_r + P_{nr}}. \quad (12.19)$$

Вероятность P_{nr} зависит от температуры

$$P_{nr} = P_{nr0} e^{-E^*/kT}. \quad (12.20)$$

В этом выражении E^* — энергия активации; P_{nr0} — коэффициент, не зависящий от температуры. Следовательно, температурная зависимость квантового выхода излучения имеет вид:

$$\eta = \frac{1}{1 + Ce^{-E^*/kT}}, \quad (12.21)$$

где $C = P_{nr0}/P_r$ — константа.

Тушение мономолекулярного свечения твердых тел, когда поглощение и излучение света происходят в одном центре, можно объяснить следующим образом. При поглощении кванта света электрон центра, переходя на возбужденный уровень, может при высокой температуре попасть на колебательный энергетический уровень, соответствующий ближайшему расположению потенциальных кривых U_0 и U_B (точка a на рис. 12.1). Затем он безызлучательным способом перейдет в основное состояние. Этот переход на рис. 12.1 показан пунктирной стрелкой. Для такого механизма температурного тушения выход люминесценции будет описываться формулой (12.21). В этом случае E^* есть энергия активации, необходимая для перевода электрона в возбужденном состоянии в точку ближайшего расположения потенциальных кривых U_0 и U_B .

Температурное тушение люминесценции полупроводника, обусловленной рекомбинацией свободных носителей заряда через локализованный центр, также будет описываться соотношением (12.21). Например, центр свечения расположен на некотором расстоянии от потолка валентной зоны (рис. 12.11, б). Такой центр, поглощая квант света, может возбуждаться (см. рис. 11.31, з). Затем ионизованный центр захватывает электрон из валентной зоны. Теперь электрон, сорванный с центра возбуждающим светом, не может излучательным путем возвратиться на центр, а вынужден некоторое время диффундировать в зоне проводимости, пока не прорекомбинирует безызлучательно с дыркой валентной зоны. Например, безызлучательные переходы при участии меди в германии имеют термическую энергию активации $E^* = 0,14$ эВ.

В случае излучательной рекомбинации свободных электронов и свободных дырок температурное тушение люминесценции можно описать формулой

$$\eta = \frac{1}{1 + CT^g}, \quad (12.22)$$

где C и g — константы.

При выводе этого соотношения использовалась теория прямой безызлучательной рекомбинации электрона зоны проводимости и дырки валентной зоны.

12.7. СПОНТАННОЕ И ВЫНУЖДЕННОЕ ИЗЛУЧЕНИЕ АТОМА

Прежде чем переходить к описанию условий возникновения лазерного излучения в полупроводниках, рассмотрим процесс поглощения и излучения света в атоме. Из всей совокупности кванто-

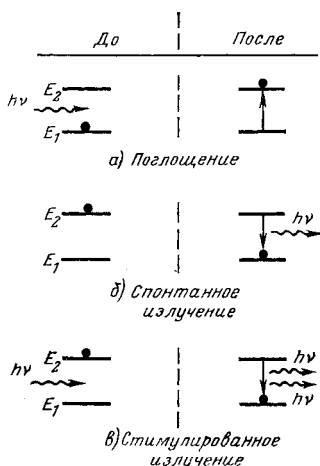


Рис. 12.16. Три основных типа переходов между энергетическими уровнями E_1 и E_2 : Черные точки обозначают состояние атома. Начальному состоянию соответствует рисунок слева; конечное состояние после протекания соответствующего процесса изображено справа

вольно независимо от воздействия внешнего поля (рис. 12.16, б). Такой переход называют спонтанным, а свечение носит название спонтанного излучения.

Спонтанные переходы являются случайными, их вероятность обозначим через A_{21} . Если N_2 — число атомов, находящихся в возбужденном состоянии, то число спонтанных переходов в единицу времени в единице объема

$$dN'_{21} = A_{21}N_2. \quad (12.24)$$

Следовательно, уменьшение количества возбужденных атомов в результате спонтанных переходов за время dt составит величину

$$-dN_2 = A_{21}N_2 dt. \quad (12.25)$$

Интегрируя это уравнение, получаем закон убывания числа возбужденных атомов со временем при спонтанных переходах:

$$N_2 = N_{20}e^{-A_{21}t}, \quad (12.26)$$

где N_{20} — число атомов, находящихся в возбужденном состоянии в момент времени $t = 0$.

Так как за каждый спонтанный переход излучается энергия $E_2 - E_1 = h\nu$, то в результате dN'_{21} переходов в единицу времени выделяется энергия в количестве

$$h\nu dN'_{21} = h\nu A_{21}N_2. \quad (12.27)$$

вых состояний, в которых может находиться атом, для простоты рассмотрения выберем только два состояния с энергией E_1 и E_2 (рис. 12.16).

Пусть на такой атом падает излучение, имеющее спектральную плотность энергии u_ν . Невозбужденный атом, поглощая энергию излучения $h\nu$, перейдет в возбужденное состояние (рис. 12.16, а). При этом переходы с поглощением совершаются тем чаще, чем больше величина u_ν . Если N_1 — число атомов в невозбужденном состоянии, а $u_\nu B_{12}$ — вероятность перехода с поглощением, то число переходов с поглощением в единицу времени в единице объема равно:

$$dN_{12} = B_{12}u_\nu N_1. \quad (12.23)$$

Атом, находясь в возбужденном состоянии в результате поглощения кванта света, через некоторый промежуток времени из возбужденного состояния возвратится в нормальное, отдавая избыток энергии $E_2 - E_1$ в виде излучения $h\nu$. Переход с излучением может произойти самопроиз-

Тогда интенсивность спонтанного свечения (энергия, испускаемая в единицу времени) равна:

$$I = h\nu A_{21} N_{20} e^{-A_{21}t} = I_0 e^{-A_{21}t}, \quad (12.28)$$

где

$$I_0 = h\nu A_{21} N_{20}. \quad (12.29)$$

Так как $A_{21}N_2dt$ есть число атомов, совершивших за время dt переход $2 \rightarrow 1$, то оно определяет количество атомов, «проживших» в возбужденном состоянии в течение времени t . Поэтому продолжительность пребывания в возбужденном состоянии всех атомов, испытывавших спонтанные переходы в течение времени от 0 до ∞ , составит

$$T = \int_0^{\infty} t A_{21} N_2 dt = \int_0^{\infty} N_{20} A_{21} t e^{-A_{21}t} dt = \frac{N_{20}}{A_{21}}. \quad (12.30)$$

Отсюда следует, что средняя продолжительность жизни атома в возбужденном состоянии

$$\tau = \frac{T}{N_{20}} = \frac{1}{A_{21}}. \quad (12.31)$$

В соответствии с этим формула (12.28) запишется в виде

$$I = I_0 e^{-t/\tau}. \quad (12.32)$$

Таким образом, *после прекращения возбуждения интенсивность спонтанного излучения во времени уменьшается по экспоненциальному закону с постоянной времени τ .*

Поскольку спонтанные переходы являются случайными, то спонтанное излучение имеет статистический характер, оно некогерентно, т. е. генерируются фотоны, распространяющиеся в произвольных направлениях и имеющие произвольные фазы, хотя их частота фиксирована. Это обусловлено тем, что акты спонтанного излучения в силу случайного характера происходят независимо друг от друга в разные моменты времени.

Однако атом, находящийся в возбужденном состоянии (на уровне E_2), под действием поля излучения может возвратиться в нормальное состояние (на уровень E_1). Такой переход называют *вынужденным*, а излучение носит название *вынужденного*, *индуцированного* или *стимулированного* (рис. 12.16, в). Индуцированные переходы будут совершаться тем чаще, чем больше плотность энергии возбуждающего света u_ν . Поэтому вероятность этих переходов в единицу времени будет $u_\nu B_{21}$. И если N_2 — число атомов в возбужденном состоянии, то число индуцированных переходов в единицу времени в единице объема составит:

$$dN_{21}' = u_\nu B_{21} N_2. \quad (12.33)$$

Следовательно, *внешнее излучение с частотой ν вызывает не только поглощение света, но и добавочное, индуцированное излучение.*

Таким образом, *среднее число переходов с излучением в единицу времени в единице объема определяется спонтанными и вынужденными переходами и равно:*

$$dN_{21} = dN'_{21} + dN''_{21} = (A_{21} + u_{\nu}B_{21})N_2. \quad (12.34)$$

В стационарном состоянии на основании принципа детального равновесия число прямых и обратных процессов в единицу времени одинаково. Следовательно,

$$dN_{12} = dN_{21} \quad (12.35)$$

или

$$(A_{21} + u_{\nu}B_{21})N_2 = u_{\nu}B_{12}N_1. \quad (12.36)$$

Отсюда найдем, что

$$\frac{N_2}{N_1} = \frac{u_{\nu}B_{12}}{A_{21} + u_{\nu}B_{11}}. \quad (12.37)$$

Поскольку равновесие имеет статистический характер, то отношение N_2/N_1 можно выразить при помощи распределения Больцмана, по которому для системы, имеющей состояния с энергией $E_1, E_2, \dots, E_m, \dots$, вероятность пребывания ее в энергетическом состоянии E_m равна:

$$W_m = Cg_me^{-E_m/kT}, \quad (12.38)$$

где g_m — статистический вес, учитывающий вырождение состояния m . Поэтому можно записать:

$$\frac{N_2}{N_1} = \frac{g_2e^{-E_2/kT}}{g_1e^{-E_1/kT}} = \frac{g_2}{g_1} e^{-(E_2-E_1)/kT}. \quad (12.39)$$

Из сравнения (12.37) и (12.39) следует, что

$$\frac{u_{\nu}B_{12}}{A_{21} + u_{\nu}B_{21}} = \frac{g_2}{g_1} e^{-(E_2-E_1)/kT}, \quad (12.40)$$

откуда с учетом $E_2 - E_1 = h\nu$ получим:

$$u_{\nu} = \frac{g_2A_{21}}{g_1B_{12}e^{h\nu/kT} - g_2B_{21}}. \quad (12.41)$$

Для определения коэффициентов A и B используем следующие предельные условия. При $T \rightarrow \infty$ плотность энергии u_{ν} должна стремиться к бесконечности, а это значит, что знаменатель в выражении (12.41) стремится к нулю. Из этого следует, что

$$g_1B_{12} = g_2B_{21}. \quad (12.42)$$

С учетом этого формула (12.41) запишется в виде

$$u_{\nu} = \frac{g_2A_{21}}{g_1B_{12}} \frac{1}{e^{h\nu/kT} - 1}. \quad (12.43)$$

При $h\nu/kT \ll 1$ в первом приближении имеем $e^{h\nu/kT} = 1 + h\nu/kT$. Следовательно,

$$u_{\nu} = \frac{g_2A_{21}kT}{g_1B_{12}h\nu}. \quad (12.44)$$

При условии $h\nu/kT \ll 1$ справедлива формула Рэлея — Джинса, по которой

$$u_\nu = \frac{8\pi\nu^2}{c^3} kT. \quad (12.45)$$

Из сопоставления (12.44) и (12.45) имеем:

$$\frac{g_2 A_{21}}{g_1 B_{12}} = \frac{8\pi h\nu^3}{c^3}, \quad (12.46)$$

откуда на основании (12.42)

$$A_{21} = \frac{8\pi h\nu^3}{c^3} \frac{g_1}{g_2} B_{12} = \frac{8\pi h\nu^3}{c^3} B_{21}. \quad (12.47)$$

Таким образом, *при вынужденном переходе генерируется фотон, имеющий ту же частицу, направление и фазу, что и фотон, индуцирующий рекомбинацию.*

Рассмотрим теперь, при каких условиях может наблюдаться индуцированное излучение. Пусть имеем параллельный пучок монохроматического света интенсивностью I . При прохождении слоя вещества толщиной Δx интенсивность света уменьшится на величину $-\Delta I = \alpha I \Delta x$. Это уменьшение интенсивности света определяется тем, что в атомах происходят переходы с поглощением энергии в количестве $u_\nu B_{12} h\nu N_1 \Delta x$ и индуцированные переходы с излучением энергии, равной $u_\nu B_{21} h\nu N_2 \Delta x$. В результате можно записать:

$$-\Delta I = \alpha I \Delta x = (B_{12} N_1 - B_{21} N_2) u_\nu h\nu \Delta x. \quad (12.48)$$

Из (12.48) следует, что поглощение и индуцированное излучение входят в нее равноправно, но с разными знаками. Если первый член выражает собой обычное поглощение, рассматриваемое как «положительное» поглощение, то индуцированное излучение можно считать «отрицательным» поглощением.

Если учесть, что плотность излучения u_ν в параллельном пучке света сечением 1 см^2 , распространяющегося со скоростью c , равна:

$$u_\nu = 1/c, \quad (12.49)$$

то коэффициент поглощения α в соответствии с равенствами (12.42), (12.48) и (12.49) приобретает вид:

$$\alpha = \frac{h\nu}{c} (B_{12} N_1 - B_{21} N_2) = \frac{h\nu}{c} \left(1 - \frac{g_1}{g_2} \frac{N_2}{N_1} \right) B_{12} N_1. \quad (12.50)$$

Учитывая (12.39), имеем:

$$\alpha = B_{12} N_1 (1 - e^{-(E_2 - E_1)/kT}) \frac{h\nu}{c}. \quad (12.51)$$

Так как $E_2 > E_1$, то экспоненциальный член в (12.51) всегда меньше единицы и величина, стоящая в скобках, положительна. Это означает, что *в случае термодинамического равновесия при любой конечной температуре суммарный коэффициент поглощения всегда положителен, т. е. $\alpha > 0$. В этом случае имеет место обычное по-*

жительное поглощение света, а индуцированные переходы лишь уменьшают его. Причиной этому является то, что в равновесном состоянии число невозбужденных атомов всегда больше числа возбужденных атомов, т. е. $N_1 > N_2$.

Из распределения Больцмана (12.39) следует, что

$$T = \frac{E_2 - E_1}{k \ln \frac{g_2 N_1}{g_1 N_2}}. \quad (12.52)$$

Поскольку $E_2 > E_1$, а в равновесии $g_1 N_1 < g_2 N_2$, то $T > 0$. Однако если выполняется неравенство

$$g_1 N_2 > g_2 N_1 \quad \text{или} \quad \frac{N_2}{N_1} > \frac{g_2}{g_1}, \quad (12.53)$$

то на основании (12.50) $\alpha < 0$. Таким образом, для того чтобы создать среду с отрицательным поглощением, необходимо осуществить неравновесное состояние, при котором число возбужденных атомов было бы больше числа атомов, находящихся в нормальном, невозбужденном состоянии. Заселенность энергетических уровней атомов, удовлетворяющих неравенству (12.53), носит название инверсной заселенности.

12.8. СТИМУЛИРОВАННОЕ ИЗЛУЧЕНИЕ ТВЕРДЫХ ТЕЛ

Одним из способов получения среды с отрицательным поглощением может явиться создание косвенным путем избытка атомов на более высоком уровне энергии по отношению к более низкому. Такой случай можно осуществить на атомной или молекулярной системе, которая может находиться в трех энергетических состояниях: в нормальном с энергией E_0 и двух возбужденных (1 и 2) с энергиями E_1 и E_2 , между которыми возможны спонтанные переходы с вероятностями A_{20} , A_{21} и A_{10} (рис. 12.17). Будем считать, что эти уровни не имеют вырождения.

В такой системе с помощью поглощаемых квантов света производится заселение возбужденного уровня 2. Селективное заселение уровня получило название оптической накачки. Уровень 1 заселяется только в результате спонтанного перехода $2 \rightarrow 1$, и число таких переходов в единицу времени равно $A_{21}N_2$. Уменьшение заселенности уровня 1 происходит только за счет переходов $1 \rightarrow 0$, и число таких переходов в единицу времени будет $A_{10}N_1$. В стационарных условиях количество заселяющих переходов равно числу обратных переходов:

$$A_{21}N_2 = A_{10}N_1. \quad (12.54)$$

Для получения индуцированного излучения необходима инверсная заселенность уровней, т. е. заселенность уровня 2 должна быть больше, чем уровня 1. Для рассматриваемой системы инверсная заселенность уровней возникнет в том случае, если

$$A_{21} < A_{10}. \quad (12.55)$$

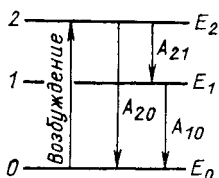


Рис. 12.17. Трехуровневая схема переходов

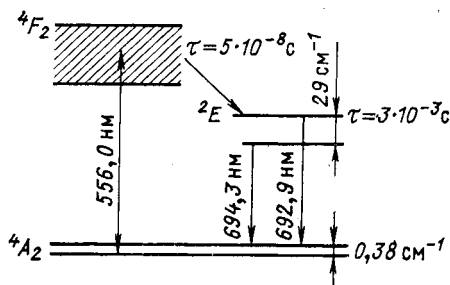


Рис. 12.18. Схема уровней в кристалле рубина

Из этого соотношения следует, что уровень 2 за счет перехода $2 \rightarrow 1$ должен опустошаться медленнее, чем уровень 1 за счет переходов $1 \rightarrow 0$. При таком соотношении вероятностей в состоянии 2 будет накапливаться больше атомов, чем в состоянии 1. Однако для осуществления такого процесса накопления атомов необходимо также, чтобы уровень 2 достаточно медленно опустошался за счет переходов $2 \rightarrow 0$, т. е. вероятность перехода из состояния 2 в состояние 0 должна быть мала.

Генераторы стимулированного излучения называют оптическими квантовыми генераторами или лазерами. В качестве примера твердотельного лазера, работающего на оптической накачке, рассмотрим рубиновый лазер. Источником света в этом лазере является кристалл рубина, который представляет собой окись алюминия Al_2O_3 , содержащий в качестве примеси замещения трехвалентные ионы хрома Cr^{3+} . У трехвалентного иона хрома, энергетическая схема которого изображена на рис. 12.18, имеется энергетическая полоса 4F_2 , а нормальное 4A_2 и возбужденное 2E состояния представляют собой узкие дублеты.

Переходу ${}^4A_2 \rightarrow {}^4F_2$ соответствует полоса поглощения с $\lambda = 556$ нм. Для состояния ${}^4F_2 \rightarrow {}^2E$ среднее время жизни $\tau = 5 \cdot 10^{-8}$ с, поэтому вероятность перехода ${}^4F_2 \rightarrow {}^2E$ велика, а для перехода ${}^2E \rightarrow {}^4A_2$ мала, так как среднее время жизни для состояния 2E примерно $5 \cdot 10^{-3}$ с.

При помощи оптической накачки производится инверсная заселенность метастабильного уровня 2E по отношению к уровню 4A_2 . Переходы ${}^2E \rightarrow {}^4A_2$ сопровождаются двумя линиями излучения с $\lambda = 694,3$ нм и $\lambda = 692,9$ нм. В спонтанном свечении эти линии слабые, так как вероятность соответствующих переходов мала.

Для создания излучения за счет вынужденных переходов кристалл рубина используется в виде стерженька длиной L в несколько сантиметров со строго параллельными и оптически отполированными торцами. Торцы стержня покрываются тонким слоем металла, например серебра. При этом один из них делается полупрозрачным.

Возникшее спонтанное излучение за счет перехода ${}^2E \rightarrow {}^4A_2$ в среде с отрицательным коэффициентом поглощения усиливается

индуцированными переходами. Его интенсивность при прохождении длины стержня $\mathcal{L}$ возрастает по закону

$$I = I_0 e^{\alpha \mathcal{L}}. \quad (12.56)$$

Увеличение интенсивности излучения будет происходить при многократном прохождении световой волны вдоль кристалла в результате отражения излучения от торцов кристалла. Многократно отражаться будут лишь те лучи, которые падают на торцевые поверхности кристалла перпендикулярно. В силу этого излучение, вызванное индуцированными переходами, сосредоточено в пределах малого телесного угла $\Delta\Omega$. В результате возникает узконаправленный пучок излучения большой интенсивности, так как яркость его возрастает в отношении $4\pi/\Delta\Omega$. Спонтанное же излучение распределяется равномерно во всех направлениях в пределах телесного угла 4π , поэтому оно уходит из кристалла через боковые поверхности.

Индукцированное излучение является когерентным. Оно происходит в том же направлении и имеет ту же фазу, что и первичное.

Оптическая накачка в рубиновом лазере может осуществляться с помощью лампы-вспышки (ксеноновой). В этом случае накопленная энергия в возбужденном состоянии 2E во время импульсной накачки высвечивается в виде вспышки. При работе рубинового лазера в непрерывном режиме требуются большая мощность оптической накачки и охлаждение кристалла. Возможно и уменьшение времени лазерной вспышки. Это достигается при помощи зеркала, которое отсутствует во время накачки, а затем быстро вводится в систему. В результате возникает лавинообразный процесс индуцированного излучения.

Создать среду с отрицательным поглощением также возможно, если искусственно понизить заселенность более низкого уровня по отношению к более высокому. Такой механизм возбуждения индуцированного свечения можно осуществить в полупроводниках, у которых экстремумы зон имеют одинаковые значения k , например InSb, InAs, InP, GaAs. У них имеет место преимущественно излучательная рекомбинация. В этом случае при прохождении тока через p - n переход в пропускном направлении рекомбинация инжектированных неосновных носителей заряда сопровождается свечением. Это свечение является спонтанным.

На излучающих диодах можно получить и стимулированное излучение, но для этого необходимо создать инверсную заселенность уровней, при которой верхние уровни должны быть заселены больше чем наполовину по отношению к нижним. Следовательно, в случае p - n перехода носители заряда должны находиться в вырожденном состоянии. Диаграмма энергетических зон инжекционного лазера, которая представлена на рис. 12.19, подобна диаграмме туннельного диода. Но степень легирования областей лазера ниже, чем туннельного диода, поэтому на вольт-амперной характеристике лазерного диода не имеется области отрицательного сопротивления или эта область пренебрежимо мала. В условиях равновесия квазиуровни

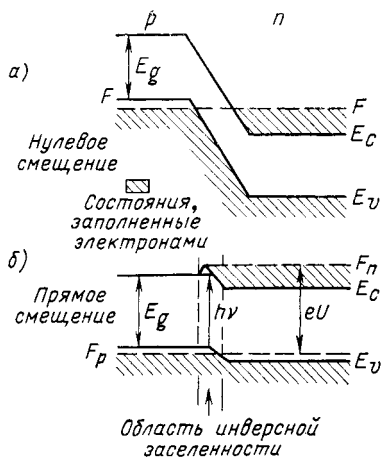


Рис. 12.19. Сильно легированный p - n переход в равновесном состоянии (а) и при прямом смещении (б)

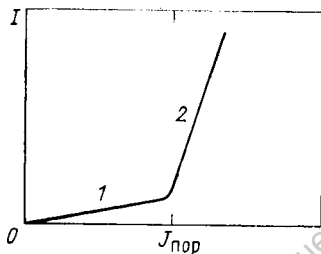


Рис. 12.20. Интенсивность излучения полупроводникового лазера в зависимости от уровня возбуждения
1 — спонтанное излучение; 2 — вынужденное излучение

Ферми лежат в зоне проводимости и валентной зоне (рис. 12.19, а). У сильно легированного p - n перехода достаточно высокое смещение в пропускном направлении может создать область инверсной заселенности (рис. 12.19, б), которая обычно очень тонкая (около 1 мкм). В этом случае электроны могут перемещаться через барьер n -области и переходить на свободные состояния валентной зоны с испусканием фотонов, энергия которых примерно равна E_g . Переходы могут также происходить между краем зоны проводимости и акцепторным уровнем (или зоной) вблизи потолка валентной зоны. Кроме того, дырки могут переходить в n -область и рекомбинировать с электронами с испусканием фотонов. При этом вид преобладающего процесса определяется относительной концентрацией примеси, значениями подвижности и времени жизни неравновесных носителей заряда.

При малых токах, текущих через p - n переход, возникает спонтанное излучение, распространяющееся по всем направлениям. Возрастание тока увеличивает скорость рекомбинации, что соответственно увеличивает плотность фотонов. Они индуцируют дальнейшую рекомбинацию. Поскольку на максимум спектрального распределения спонтанного излучения приходится больше фотонов, на этой частоте вызывается больше переходов, чем на краях спектра. Это приводит к значительному сужению спектра излучения и, следовательно, усилению интенсивности излучения в его максимуме.

Типичный характер зависимости интенсивности излучения лазера от уровня возбуждения представлен на рис. 12.20. Как видно, ниже некоторого порогового уровня возбуждения $J_{пор}$ излучение в основном имеет спонтанный характер (кривая 1). При возрастании тока усиление также возрастает. Но когда достигнут порог генерации, интенсивность вынужденного излучения резко увеличивается (кри-

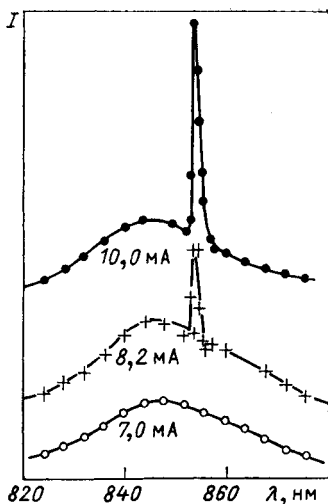


Рис. 12.21. Спектр GaAs-лазера при токах ниже и выше пороговых

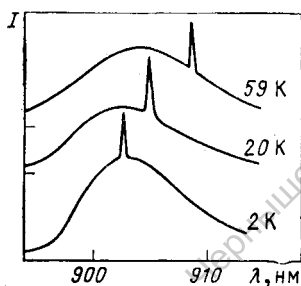


Рис. 12.22. Зависимость излучения GaAs-лазерного диода от температуры

вая 2) и из окружающих плоскостей p - n перехода будет выходить когерентное излучение.

На рис. 12.21 изображен спектр излучения типичного лазерного диода. При значениях тока ниже пороговых спектр излучения определяется спонтанными переходами и представляет собой относительно широкую полосу. При возрастании тока выше пороговых значений спектр излучения значительно сужается. Зависимость спектрального излучения этого лазера от температуры представлена на рис. 12.22. Как видно, при возрастании температуры линия вынужденного излучения сдвигается в область более длинных волн, что согласуется с температурным изменением ширины запрещенной зоны.

Инверсная заселенность в полупроводниках может быть осуществлена также фотовозбуждением и возбуждением с помощью пучка электронов. При оптическом возбуждении инверсия осуществляется фотонами, обладающими энергией, превышающей ширину запрещенной зоны. Пусть при абсолютном нуле за счет фотовозбуждения состояния зоны проводимости полупроводника с прямыми оптическими переходами заполнены до квазиуровня Ферми ξ_n и введены дырки в валентную зону до квазиуровня Ферми ξ_p , как это изображено на рис. 12.23, а. При такой инверсной заселенности ни одно состояние, участвующее в процессе излучения, не может участвовать

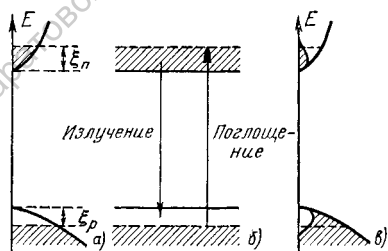


Рис. 12.23. Заполнение зон до вырождения, предотвращающее межзонное поглощение на частоте излучения

в поглощении (рис. 12.23, б). При конечных температурах распределение носителей заряда будет такое, как показано на рис. 12.23, в. Хотя для всей системы тепловое равновесие отсутствует, но носители заряда в данной энергетической зоне находятся в тепловом равновесии друг с другом. Фотоны, имеющие энергию $h\nu$ [$E_g < h\nu < (F_n - F_p)$], вызывают переходы электронов на нижние уровни, что приводит в свою очередь к возникновению стимулированного излучения.

Получения вырожденной электронно-дырочной заселенности в полупроводниках можно достичь за счет перехода электронов из валентной зоны в зону проводимости при бомбардировке кристалла пучком электронов. Преимуществом оптического возбуждения и возбуждения с помощью электронного пучка является то, что в этих случаях возбуждается большой объем материала и не требуется создание p - n переходов. Это особенно важно для материалов, в которых трудно получить p - n переход.

ГЛАВА ТРИНАДЦАТАЯ

ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

13.1. ВНУТРЕННИЙ ФОТОЭФФЕКТ

Изменение электрического сопротивления полупроводника под действием излучения называют фоторезистивным эффектом. Добавочная проводимость, обусловленная носителями заряда, созданными оптической генерацией, носит название фотопроводимости.

При внутреннем фотоэффекте первичным актом является поглощение фотона. Поэтому процесс образования свободных носителей заряда будет происходить по-разному в зависимости от особенностей процесса поглощения света полупроводником. При межзонных переходах имеет место собственная фотопроводимость (переход 1, рис. 13.1). Для полупроводников с прямыми зонами при вертикальных переходах энергия фотона $h\nu$ должна быть не меньше ширины запрещенной зоны, т. е.

$$h\nu \geq E_g. \quad (13.1)$$

В случае непрямых переходов, когда сохранение квазиимпульса обеспечивается за счет излучения фотона с энергией E_p , длинноволновый край спектра фотопроводимости будет лежать при

$$h\nu = E_g + E_p. \quad (13.2)$$

Для сильно легированного полупроводника n -типа, когда уровень Ферми рас-

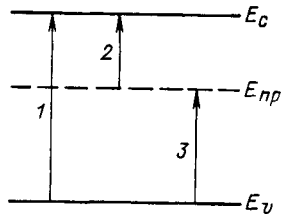


Рис. 13.1. Схема возможных оптических переходов, обуславливающих фотопроводимость

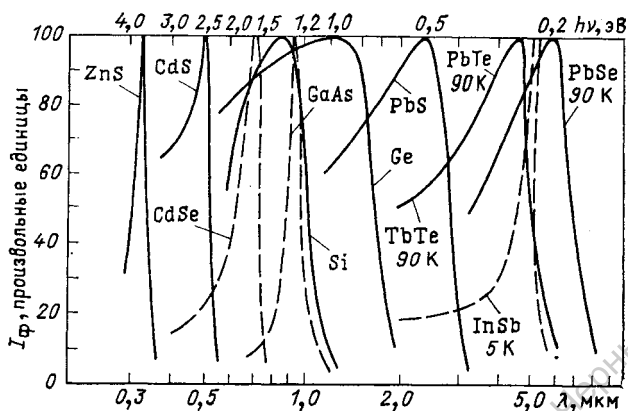


Рис. 13.2. Спектральное распределение фототока некоторых полупроводников в области собственного поглощения

положен выше дна зоны проводимости на величину ξ_n , длинноволновая граница спектра фототока будет соответствовать

$$h\nu = E_g + \xi_n. \quad (13.3)$$

В сильно легированном полупроводнике p -типа уровень Ферми лежит на величину ξ_p ниже края валентной зоны, поэтому

$$h\nu = E_g + \xi_p. \quad (13.4)$$

Собственная полоса поглощения, всегда имеющая отчетливо выраженную длинноволновую границу, в принципе может иметь и коротковолновую. Однако у многих полупроводников зона проводимости перекрывается вышележащими разрешенными зонами, образуя сплошную зону. Поэтому спектральное распределение фоточувствительности в зависимости от энергии фотонов или длины волны света должно простираться далеко в коротковолновую область. Но с увеличением энергии фотонов увеличивается коэффициент собственного поглощения, а следовательно, будет иметь место и увеличение фотопроводимости. И если квантовая эффективность остается постоянной, то при больших энергиях область поглощения и, следовательно, область генерации фотоносителей располагается вблизи поверхности полупроводника. В приповерхностной области полупроводника время жизни носителей заряда меньше, чем в объеме образца. Изменение времени жизни неравновесных носителей заряда приведет к уменьшению фотопроводимости в области коротких длин волн (рис. 13.2).

При наличии в запрещенной зоне полупроводника локальных уровней оптическое поглощение может вызвать переходы электронов между уровнями примеси и энергетическими зонами (переходы 2 и 3, рис. 13.1). В этом случае фотопроводимость называют *примесной фотопроводимостью*. Поскольку энергия ионизации примеси E_{up} меньше ширины запрещенной зоны E_g , то спектр

примесного фототока расположен в длинноволновой области по отношению к спектру собственной фотопроводимости. В качестве примера на рис. 13.3 приведен спектр фототока германия, легированного медью и цинком.

При экситонном поглощении света имеет место создание связанной пары электрон—дырка, которая является электрически нейтральным образованием. Поэтому экситонное поглощение первоначально не ведет к возникновению свободных носителей заряда. Однако в реальных кристаллических структурах экситоны могут диссоциировать при взаимодействии с фононами, примесными центрами и дефектами решетки. Таким образом, образование экситонов в конечном итоге ведет к возникновению свободных носителей заряда, а следовательно, фототока. При низких температурах экситонное поглощение, характеризующееся узкими полосами поглощения, определяет и узкие полосы фототока. При этом спектр фототока в области экситонного поглощения будет зависеть как от состояния поверхности, так и от соотношения диффузионной длины экситонов $L_{\text{экс}}$ или амбиполярной диффузионной длины L . Так, например, при обедненной поверхности, когда время жизни неосновных носителей заряда в приповерхностной области полупроводника τ_s будет меньше времени жизни τ , характерного для его объема, максимумы полос экситонного поглощения будут соответствовать минимумам на кривой фототока, если $L_{\text{экс}} < L$, максимумам, если $L_{\text{экс}} > L$. Напротив, если $\tau_s > \tau$, то следует ожидать противоположных соотношений между положением максимумов полос экситонного поглощения и фототока в зависимости от соотношения $L_{\text{экс}}$ и L .

Состояние поверхности можно легко изменить полупроводника можно легко изменить путем различных воздействий на нее (механическое, химическое

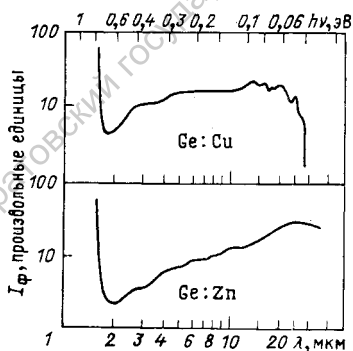


Рис. 13.3. Спектр фототока германия, легированного медью и цинком

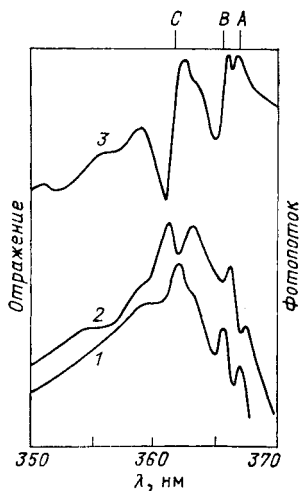


Рис. 13.4. Спектры отражения (кривая 3) и фототока монокристалла окиси цинка до (кривая 2) и после травления (кривая 1)

и т. д.). Таким образом можно изменить характер наблюдаемого спектра фототока, обусловленного экситонным поглощением. Для примера на рис. 13.4 представлены спектры фототока монокристалла ZnO. Положение максимумов экситонных полос поглощения A , B и C (показано стрелками) определено из спектра отражения (кривая 3). Видно, что максимумы полос экситонного поглощения соответствуют минимумам на кривой фототока (кривая 2). Травление поверхности такого кристалла приводит к случаю, когда максимумам поглощения соответствуют максимумы фототока (кривая 1).

Поглощение света свободными носителями заряда и колебаниями решетки непосредственно не могут привести к изменению концентрации носителей заряда. Однако возрастание концентрации носителей заряда в этих случаях может происходить в результате вторичных эффектов, когда поглощение света значительно увеличивает кинетическую энергию свободных носителей заряда, которые затем отдают свою энергию на возбуждение носителей заряда.

13.2. ФОТОПРОВОДИМОСТЬ

Избыточные электроны Δn и дырки Δp , созданные оптической генерацией, могут иметь энергии, значительно большие, чем средняя энергия равновесных носителей заряда. Однако в результате взаимодействия с фононами и дефектами кристаллической решетки энергия неравновесных носителей заряда за время около 10^{-10} — 10^{-12} с приобретает такое распределение по энергиям и квазимпульсам, как у равновесных носителей. Поэтому подвижность неравновесных носителей заряда не будет отличаться от подвижности равновесных носителей. Следовательно, полная проводимость полупроводника определяется равновесными носителями заряда n_0 , p_0 и фотоносителями Δn , Δp и равна:

$$\sigma = e[(n_0 + \Delta n)\mu_n + (p_0 + \Delta p)\mu_p]. \quad (13.5)$$

Так как темновая проводимость $\sigma_T = \sigma_0 = e(n_0\mu_n + p_0\mu_p)$, то фотопроводимость полупроводника, обусловленная непосредственным действием излучения, есть

$$\sigma_{\phi} = \sigma - \sigma_0 = e(\Delta n\mu_n + \Delta p\mu_p). \quad (13.6)$$

Естественно, что концентрации избыточных носителей заряда Δn и Δp зависят от интенсивности и длины волны света. Пусть на слой вещества толщиной dx , имеющего коэффициент поглощения α , падает свет интенсивности I . Тогда количество световой энергии поглощаемой в единицу времени в единице объема этого вещества

$$\frac{dI}{dx} = \alpha I. \quad (13.7)$$

Следовательно, при поглощении квантов света энергии $h\nu$ в единице объема полупроводника в единицу времени для области собственного поглощения образуются избыточные электроны и дырки в количестве

$$G = \beta \alpha I / h\nu. \quad (13.8)$$

Здесь коэффициент пропорциональности β , который обычно называют квантовым выходом фотоионизации, определяет число пар носителей заряда (или число носителей заряда при примесной фотопроводимости), образуемых одним поглощенным фотоном, если интенсивность света измерять числом квантов в секунду на единицу поверхности.

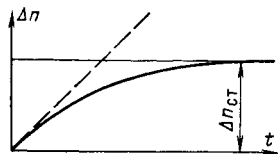


Рис. 13.5. Изменение во времени концентрации носителей заряда, возбужденных светом

Однако сразу после начала освещения фотопроводимость полупроводника не достигает максимального значения, ибо по мере увеличения концентрации неравновесных носителей заряда нарастает процесс рекомбинации. Поскольку скорость генерации неравновесных носителей остается постоянной при неизменной интенсивности света, то через какой-то промежуток времени интенсивность рекомбинации достигнет интенсивности генерации и установится стационарное состояние, характеризующееся постоянным значением концентрации фотоносителей заряда $\Delta n_{ст}$ и $\Delta p_{ст}$ (рис. 13.5).

Стационарные концентрации избыточных носителей заряда можно определить, если воспользоваться уравнением непрерывности, в котором генерационный член записан в виде (13.8) в предположении однородной генерации. Поэтому

$$\Delta n_{ст} = G_n \tau_n = \beta \alpha I \tau_n / h\nu; \quad (13.9)$$

$$\Delta p_{ст} = G_p \tau_p = \beta \alpha I \tau_p / h\nu, \quad (13.10)$$

а стационарная фотопроводимость равна:

$$\sigma_{ф,ст} = e\beta\alpha(\mu_n\tau_n + \mu_p\tau_p) \frac{I}{h\nu}. \quad (13.11)$$

Отношение фотопроводимости $\sigma_{ф}$ к интенсивности света I определяет удельную фоточувствительность полупроводника

$$S_{ф} = \sigma_{ф} / I. \quad (13.12)$$

Если один из членов в скобках соотношения (13.11) значительно больше другого (за счет разницы в значениях подвижности или времени жизни электронов и дырок), то фотопроводимость определяется носителями заряда одного знака и ее называют монополярной. В этом случае

$$\sigma_{ф,ст} = e\beta\alpha\mu \frac{I}{h\nu}. \quad (13.13)$$

Выражение для стационарного значения плотности фототока будет иметь вид:

$$J_{ф} = \sigma_{ф,ст} \mathcal{E} = e\beta\alpha(\mu_n\tau_n + \mu_p\tau_p) \frac{I\mathcal{E}}{h\nu}. \quad (13.14)$$

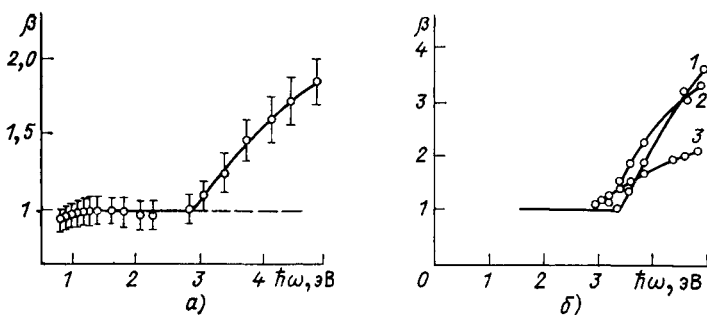


Рис. 13.6. Зависимость квантового выхода (число возбужденных электронов на один поглощенный фотон) от энергии фотона в германии $T = 300$ К (а) и кремний (б): 1 — 100 К; 2 — 300 К; 3 — 400 К

Но $\mu_n \mathcal{E} = v_n$, а $\mu_p \mathcal{E} = v_p$, и если через l обозначить размер полупроводника в направлении поля, то время дрейфа $t_n = l/v_n$ и $t_p = l/v_p$. С учетом этого (13.14) запишется как

$$J_\Phi = e\beta\alpha \left(\frac{\tau_n}{t_n} + \frac{\tau_p}{t_p} \right) \frac{Il}{h\nu}. \quad (13.15)$$

Если величины, входящие в (13.15), известны, то, измеряя J_Φ , можно определить квантовый выход β . На рис. 13.6, а приведена спектральная характеристика квантового выхода для германия и кремния. Видно, что при 300 К вплоть до 2,7 эВ для германия и вплоть до 3 эВ у кремния $\beta = 1$. При дальнейшем росте энергии фотона квантовый выход резко увеличивается. Это происходит потому, что поглощение фотона столь большой энергии сопровождается возникновением «горячих» носителей заряда, обладающих энергией, достаточной для образования вторичных электронно-дырочных пар путем ударной ионизации. Поскольку ширина запрещенной зоны кремния уменьшается при повышении температуры, то граница роста квантового выхода, как следует из рис. 13.6, б, смещается в сторону меньших энергий.

13.3. РЕЛАКСАЦИЯ ФОТОПРОВОДИМОСТИ

Пусть полупроводник освещается импульсом света прямоугольной формы, как это изображено на рис. 13.7, а. Стационарное значение фотопроводимости достигается не мгновенно, а лишь через некоторое время после начала освещения (рис. 13.7, б). При выключении света неравновесная проводимость исчезнет также через некоторое время после прекращения освещения (рис. 13.7, в). Нас будут интересовать законы нарастания и спада фотопроводимости в зависимости от уровня освещения. Рассмотрим случаи малого и большого уровня возбуждения.

Изменение концентрации избыточных носителей заряда в единицу времени $d\Delta n/dt$ есть разность между скоростями генерации

и рекомбинации носителей заряда. Поскольку скорость генерации $G = \beta \alpha I / h\nu$, то

$$\frac{d\Delta n}{dt} = \frac{\beta \alpha I}{h\nu} - R, \quad (13.16)$$

где R учитывает уменьшение концентрации избыточных носителей заряда в результате процесса рекомбинации.

При малом уровне освещения, для которого $\Delta n \ll (n_0 + p_0)$, согласно выражению (7.18) $R = \Delta n / \tau$, поэтому

$$\frac{d\Delta n(t)}{dt} = \frac{\beta \alpha I}{h\nu} - \frac{\Delta n(t)}{\tau}. \quad (13.17)$$

Разделяя переменные и интегрируя с учетом начального условия $\Delta n = 0$ при $t = 0$, получаем выражения, характеризующие изменения концентрации избыточных электронов во времени:

$$\Delta n = \beta \alpha \tau \frac{I}{h\nu} (1 - e^{-t/\tau}) = \Delta n_{\text{ст}} (1 - e^{-t/\tau}); \quad (13.18)$$

аналогично для процесса спада

$$\Delta n = \beta \alpha \tau \frac{I}{h\nu} e^{-t/\tau} = \Delta n_{\text{ст}} e^{-t/\tau}. \quad (13.19)$$

Отсюда следует, что *релаксация фотопроводимости при малой освещенности определяется экспоненциальным законом с постоянной времени, соответствующей времени жизни неравновесных носителей заряда.*

При большом уровне возбуждения, когда $\Delta n \gg (n_0 + p_0)$, согласно (7.20) $R = \gamma (\Delta n)^2$ и (13.17) примет вид

$$\frac{d\Delta n}{dt} = \beta \alpha \frac{I}{h\nu} - \gamma (\Delta n)^2. \quad (13.20)$$

Для решения этого уравнения также используем начальное условие: $\Delta n = 0$ при $t = 0$. В результате получим, что концентрация избыточных электронов при квадратичной рекомбинации описывается выражением для нарастания

$$\Delta n = \sqrt{\frac{\beta \alpha I}{\gamma h\nu}} \operatorname{th} t \sqrt{\frac{\beta \alpha \gamma I}{h\nu}}, \quad (13.21)$$

а для процесса спада

$$\Delta n = \sqrt{\frac{\beta \alpha I}{\gamma h\nu}} \frac{1}{1 + \sqrt{\frac{\beta \alpha \gamma I}{h\nu}} t}. \quad (13.22)$$

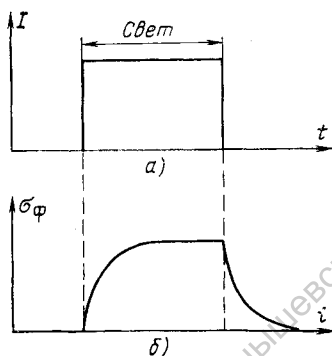


Рис. 13.7. Релаксация фотопроводимости (б) при возбуждении ее импульсом света прямоугольной формы (а)

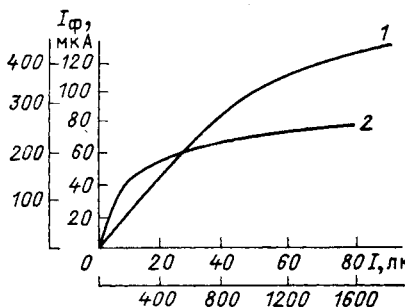


Рис. 13.8. Зависимость стационарного фототока от интенсивности освещения для сернистого висмута:

1 — слабое освещение; 2 — сильное

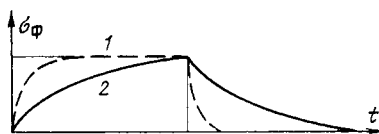


Рис. 13.9. Кривые релаксации фотопроводимости при отсутствии (1) и наличии (2) центров захвата

Таким образом, фотопроводимость зависит от интенсивности света. При линейном законе рекомбинации стационарная фотопроводимость согласно (13.19) пропорциональна интенсивности света, т. е.

$$\sigma_{\phi} \sim I/h\nu \text{ и } J_{\phi} \sim I/h\nu. \quad (13.23)$$

При квадратичной рекомбинации из (13.22) следует, что

$$\sigma_{\phi} \sim \sqrt{\frac{I}{h\nu}} \text{ и } J_{\phi} \sim \sqrt{\frac{I}{h\nu}}. \quad (13.24)$$

В общем случае можно считать, что

$$\sigma_{\phi} \sim \frac{I^b}{h\nu}. \quad (13.25)$$

При $b = 1$ фоторезистивный эффект называют линейным, при $b < 1$ — сублинейным. Следовательно, зависимость стационарного фототока полупроводника от интенсивности света должна иметь два участка; при слабой освещенности — линейный, а при большой освещенности — сублинейный (рис. 13.8).

Если в полупроводнике имеются центры захвата, то они оказывают влияние на кинетику фотопроводимости. Свободные электроны будут не только рекомбинировать с дырками, но будут также захватываться центрами прилипания, что уменьшит скорость нарастания стационарной концентрации носителей заряда. При выключении возбуждающего света опустошение ловушек будет затягивать спад концентрации носителей заряда. В результате этого будет иметь место замедление процесса нарастания и спада фотопроводимости (рис. 13.9).

13.4. ФОТОПРОВОДИМОСТЬ ПРИ НАЛИЧИИ ПОВЕРХНОСТНОЙ РЕКОМБИНАЦИИ И ДИФФУЗИИ НОСИТЕЛЕЙ ЗАРЯДА

В предыдущем анализе фотопроводимости предполагалась однородная генерация носителей заряда по всему образцу. Кроме того, не была учтена поверхностная рекомбинация, которая приводит к уменьшению концентрации неравновесных носителей заряда вблизи поверхности полупроводника.

Рассмотрим теперь случай, когда падающее излучение поглощается неравномерно по толщине образца и имеет место диффузия носителей заряда.

При отсутствии центров прилипания для вычисления стационарной концентрации избыточных носителей заряда нужно найти решение уравнения непрерывности для неосновных неравновесных носителей заряда (например, для дырок):

$$-\frac{\Delta p}{\tau_p} + D \frac{d^2 \Delta p}{dx^2} + Ge^{-\alpha x} = 0. \quad (13.26)$$

В этом уравнении первое слагаемое учитывает рекомбинацию; второе — диффузию неравновесных дырок, которая возникает как результат неоднородной генерации носителей заряда и влияния поверхностной рекомбинации; третье — генерацию носителей заряда, которая согласно (11.4) будет

$$Ge^{-\alpha x} = \frac{\beta (1 - R_s)^2 I_0}{h\nu} e^{-\alpha x}, \quad (13.27)$$

где β — квантовый выход; $h\nu$ — энергия падающего кванта света; R_s — коэффициент отражения поверхности полупроводника.

Если излучение интенсивности I_0 падает на поверхность образца, то граничное условие для $x = 0$ можно записать в виде

$$D \left. \frac{d\Delta p}{dx} \right|_{x=0} = s\Delta p \Big|_{x=0}, \quad (13.28)$$

где s — скорость поверхностной рекомбинации на освещенной поверхности.

Условия на второй поверхности образца толщиной d можно не учитывать, если считать образец достаточно толстым: $d \gg L$ и $d \gg 1/\alpha$.

Общее решение (13.26) имеет вид:

$$\Delta p(x) = Ae^{-x/L} - \frac{\tau Ge^{-\alpha x}}{L^2 \alpha^2 - 1}. \quad (13.29)$$

Используя граничное условие (13.28), будем иметь:

$$\Delta p(x) = \frac{G\tau}{\alpha^2 L^2 - 1} \left[\frac{\alpha L^2 + s\tau}{L + s\tau} e^{-x/L} - e^{-\alpha x} \right]. \quad (13.30)$$

Полное количество генерированных дырок, отнесенных к единице поверхности образца толщиной d , получим, если проинтегрируем $\Delta p(x)$ в пределах от 0 до d и учтем (13.27):

$$\begin{aligned} \Delta P &= \int_0^d \Delta p(x) dx = \frac{\beta (1 - R_s)^2 I_0 \tau \alpha}{h\nu (\alpha^2 L^2 - 1)} \times \\ &\times \left[\frac{\alpha L^2 + s\tau}{L + s\tau} L (1 - e^{-d/L}) - \frac{1}{\alpha} (1 - e^{-\alpha d}) \right]. \end{aligned} \quad (13.31)$$

Поскольку $\Delta P = \Delta N$, а фотопроводимость пропорциональна ΔP и ΔN (13.6), то (13.31) позволяет проанализировать влияние скоро-

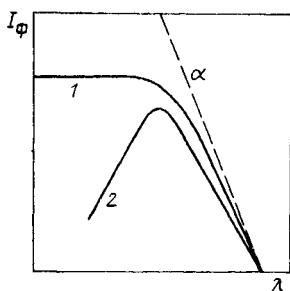


Рис. 13.10. Зависимость фототока от длины волны поглощенного света при отсутствии (1) и наличии (2) поверхностной рекомбинации

сти поверхностной рекомбинации на фотопроводимость, а также выяснить некоторые особенности спектральной зависимости фотопроводимости собственного полупроводника. Для этого кривую собственного поглощения условно разделим на два участка.

1. Область малого коэффициента поглощения, когда $\alpha d < 1$ и $\alpha L \ll 1$. В этом случае $1 - e^{-\alpha d} \approx \alpha d$, а $e^{-d/L} \approx 0$, так как $d \gg L$. Поэтому (13.31) примет вид:

$$\Delta P = \frac{\beta (1 - R_s)^2 I_0 \tau \alpha}{h\nu} \left(d - \frac{\alpha L^2 + s\tau}{L + s\tau} L \right) \approx \frac{\beta (1 - R_s)^2 I_0 \tau}{h\nu} \alpha \left(d - \frac{s\tau L}{L + s\tau} \right). \quad (13.32)$$

Из этого выражения следует, что в области малого коэффициента поглощения фотопроводимость линейно растет с увеличением α . При этом, если $s\tau \ll d$, поверхностная рекомбинация слабо влияет на фотопроводимость, если же $s\tau \gg d$, то фотопроводимость будет меньше (кривая 2, рис. 13.10), чем при $s = 0$ (кривая 1, рис. 13.10).

2. Область средних значений коэффициента поглощения, когда $\alpha d > 1$, а $\alpha L < 1$. В этом случае

$$\Delta P = \frac{\beta (1 - R_s)^2 I_0 \tau \alpha}{h\nu} \left[\frac{1}{\alpha} - \frac{\alpha L^2 + s\tau}{L + s\tau} \right] \approx \frac{\beta (1 - R_s)^2 I_0 \tau \alpha}{h\nu} \left[1 - \frac{\alpha s\tau}{L + s\tau} \right]. \quad (13.33)$$

Видно, что с дальнейшим ростом коэффициента поглощения α фотопроводимость будет убывать (кривая 2, рис. 13.10). В случае отсутствия поверхностной рекомбинации ($s = 0$) фотопроводимость стремится к насыщению (кривая 1, рис. 13.10).

Следовательно, наличие поверхностной рекомбинации приводит к появлению максимума на кривой спектральной зависимости фотопроводимости.

13.5. ЭФФЕКТ ДЕМБЕРА

Если полупроводник осветить светом с длиной волны, соответствующей области собственного поглощения, то вблизи поверхности образца, где происходит поглощение света, возникнут электронно-дырочные пары. Эти носители заряда диффундируют из освещенной области в глубь полупроводника (рис. 13.11). Поскольку электроны обладают большей подвижностью, чем дырки, то электроны продвигаются дальше в глубь кристалла, чем дырки. Эта разница в диффузии носителей заряда приведет к тому, что поверхность полупроводника окажется положительно заряженной по отношению к его объему. Возникшее электрическое поле направлено так, что оно

ускоряет дырки, имеющие меньшую подвижность, и замедляет подвижные электроны, поэтому суммарный ток равен нулю. Объемная ЭДС, возникающая в освещенном полупроводнике вследствие различия в коэффициентах диффузии электронов и дырок, получила название ЭДС Дембера.

Вычисление ЭДС Дембера проведем в предположении, что в полупроводнике мало ловушек и нет поверхностной рекомбинации. Для этого необходимо решить уравнения

$$J_n = e(n_0 + \Delta n) \mu_n \mathcal{E}_D + e D_n \frac{d\Delta n}{dx}; \quad (13.34)$$

$$J_p = e(p_0 + \Delta p) \mu_p \mathcal{E}_D - e D_p \frac{d\Delta p}{dx}, \quad (13.35)$$

где $\mathcal{E}_D$ — напряженность поля Дембера и $\Delta n = \Delta p$ — концентрация избыточных носителей заряда.

В стационарных условиях полная плотность тока равна нулю. Так как $e D_p = \mu_p kT$ и $e D_n = \mu_n kT$, то

$$J = J_n + J_p = e[(n_0 + \Delta n) \mu_n + (p_0 + \Delta p) \mu_p] \mathcal{E}_D + e \left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx} \right) = 0 \quad (13.36)$$

и для напряженности поля получим выражение

$$\mathcal{E}_D = - \frac{kT}{e} \frac{(\mu_n - \mu_p) \frac{d\Delta n}{dx}}{n\mu_n + p\mu_p}, \quad (13.37)$$

где $n = n_0 + \Delta n$ и $p = p_0 + \Delta p$.

Принтегрируем (13.37) по x от поверхности ($x = 0$) до такой точки, где концентрация неравновесных носителей обращается в нуль. Это расстояние должно быть порядка нескольких длин диффузии ($x = 3L$). Так как $\sigma = e(n\mu_n + p\mu_p)$, а $\Delta\sigma = e(\mu_n + \mu_p) \Delta n$, то

$$\mathcal{E}_D = - \frac{kT}{e} \frac{b-1}{b+1} \frac{1}{\sigma} \frac{d\Delta\sigma}{dx}, \quad (13.38)$$

где $b = \mu_n / \mu_p$.

Следовательно,

$$U_D = - \frac{kT}{e} \frac{b-1}{b+1} \int_{x=0}^{x=3L} \frac{d\Delta\sigma}{\sigma} = \frac{kT}{e} \frac{b-1}{b+1} \ln \frac{\sigma(0)}{\sigma(3L)}. \quad (13.39)$$

На расстоянии $x = 3L$ все неравновесные носители прорекомбинируют, а поэтому $\sigma(3L) = e(n_0\mu_n + p_0\mu_p)$. У поверхности полупроводника $\sigma(0) = e(n_0 + \Delta n) \mu_n + e(p_0 + \Delta p) \mu_p = \sigma(3L) + e(\mu_n + \mu_p) \Delta n$. Поэтому

$$U_D = \frac{kT}{e} \frac{b-1}{b+1} \ln \left[1 + \frac{(b+1) \Delta n}{bn_0 + p_0} \right]. \quad (13.40)$$

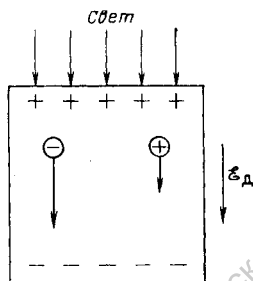


Рис. 13.11. Возникновение ЭДС Дембера

Из (13.39) следует, что ЭДС Дембера тем больше, чем сильнее различаются подвижности электронов и дырок. ЭДС Дембера обычно очень мала, она немного больше kT/e . При $\mu_n = \mu_p$ $U_D = 0$.

13.6. ФОТОЭЛЕКТРОМАГНИТНЫЙ ЭФФЕКТ

Если полупроводник, освещенный сильно поглощаемым светом, поместить в магнитное поле, перпендикулярное направлению диффузии носителей заряда, то в нем возникает поперечная ЭДС (рис. 13.12). Это явление называют фотоэлектромагнитным эффектом (ФЭМ-эффектом).

На диффундирующие электроны и дырки, созданные светом, в магнитном поле действует сила Лоренца, отклоняющая их в направлении, перпендикулярном направлению их движения и направлению магнитного поля. При этом потоки электронов и дырок отклоняются в разные стороны. Вследствие этого у противоположных сторон образца, как это представлено на рис. 13.13, скапливаются заряды противоположного знака и возникает электрическое поле, а следовательно, возникает напряжение U_y ФЭМ-эффекта. Накопление зарядов и возрастание ЭДС будут продолжаться до тех пор, пока ток проводимости, обусловленный возникшим электрическим полем, не компенсирует магнитно-диффузионный ток.

Определение U_y и тока короткого замыкания проведем для полупроводника, у которого скорость поверхностной рекомбинации мала, а поглощение света происходит в приповерхностном слое образца с образованием электронно-дырочных пар.

В случае слабого магнитного поля ($\mu B \ll 1$) углы Холла для электронов φ_n и для дырок φ_p согласно (6.88) запишутся в виде

$$\varphi_n = -\mu_n^{(H)} B; \quad (13.41)$$

$$\varphi_p = \mu_p^{(H)} B, \quad (13.42)$$

где $\mu_n^{(H)}$ и $\mu_p^{(H)}$ — холловские подвижности электронов и дырок. Они связаны с дрейфовыми подвижностями соотношениями $\mu_n^{(H)} = r\mu_n$ и $\mu_p^{(H)} = r\mu_p$, где r — холл-фактор, зависящий от механизма рассеяния (см. § 6.5).

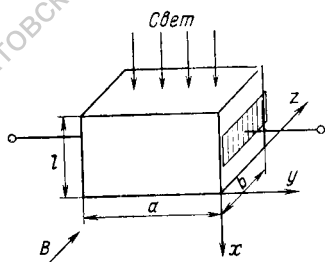


Рис. 13.12. Образец для измерения фотоэлектромагнитного эффекта

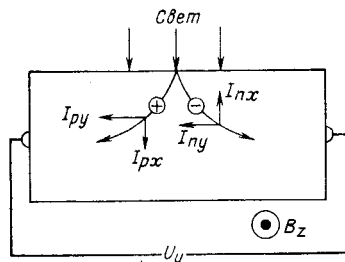


Рис. 13.13. Возникновение напряжения U_y при фотоэлектромагнитном эффекте

Если J_{nx} и J_{px} есть плотности диффузионных токов в направлении оси x , то согласно рис. 13.13 они равны

$$J_{ny} = J_{nx} \operatorname{tg} \varphi_n = -r\mu_n B J_{nx}; \quad (13.43)$$

$$J_{py} = J_{px} \operatorname{tg} \varphi_p = r\mu_n B J_{px}; \quad (13.44)$$

Плотность полного диффузионного тока составит:

$$J_y = J_{ny} + J_{py} = rB(\mu_p J_{px} - \mu_n J_{nx}). \quad (13.45)$$

Если в полупроводнике нет ловушек, то $\Delta n = \Delta p$. С учетом этого диффузионные токи запишутся в виде

$$J_{px} = -J_{nx} = -eD \frac{d\Delta n}{dx}, \quad (13.46)$$

где D — коэффициент амбиполярной диффузии.

Подставляя (13.46) в (13.45), получаем:

$$J_y = -erDB(\mu_n + \mu_p) \frac{d\Delta n}{dx}. \quad (13.47)$$

Для того чтобы найти плотность тока короткого замыкания во внешней цепи $J_{к.з.}$, необходимо подсчитать плотность суммарного диффузионного тока, текущего через всю боковую поверхность (сечение образца в плоскости xz). Если D считать величиной постоянной, а b — размером образца по оси z , то

$$\begin{aligned} J_{к.з.} &= \int_0^l J_y b dx = -erBD(\mu_n + \mu_p) b \int_0^l d\Delta n = \\ &= erDB(\mu_n + \mu_p) b [\Delta n(0) - \Delta n(l)]. \end{aligned} \quad (13.48)$$

В стационарном состоянии полный ток через все сечение образца в плоскости xz равен нулю, т. е. ток короткого замыкания равен полному току проводимости, который определяется напряжением ФЭМ-эффекта $U_{фэм}$ и полной проводимостью образца σ :

$$J = U_{фэм} \sigma = J_{к.з.} \quad (13.49)$$

Поскольку полная проводимость образца

$$\sigma = \frac{bl}{a} \sigma_0 + \frac{b}{a} \int_0^l \Delta \sigma dx = \frac{bl}{a} \sigma_0 + \frac{eb}{a} (\mu_n + \mu_p) \int_0^l \Delta n dx, \quad (13.50)$$

где a — размер образца по оси y ; σ_0 — проводимость в темноте, то полная величина фото-ЭДС составляет:

$$U_{фэм} = \frac{J_{к.з.}}{\sigma} = \frac{eraDB(\mu_n + \mu_p) [\Delta n(0) - \Delta n(l)]}{l\sigma_0 + e(\mu_n + \mu_p) \int_0^l \Delta n dx}. \quad (13.51)$$

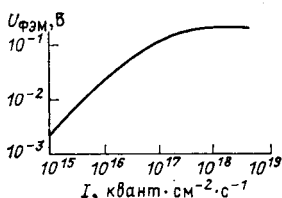


Рис. 13.14. Зависимость фотомагнитной ЭДС от интенсивности освещения для германия

Следовательно, напряжение фотоэлектромагнитного эффекта пропорционально индукции магнитного поля и зависит от концентрации неравновесных носителей заряда на противоположных сторонах образца.

Если генерация носителей заряда происходит вблизи освещенной поверхности образца, то их концентрация из-за диффузии вдоль оси x изменяется по закону

$$\Delta n = \Delta n(0) e^{-x/L}. \quad (13.52)$$

В том случае, если толщина образца l значительно превосходит диффузионную длину L , т. е. $l \gg L$, то $\Delta n(l) = 0$, а

$$\int_0^l \Delta n dx = \Delta n(0) \int_0^l e^{-x/L} dx = \Delta n(0) L. \quad (13.53)$$

Следовательно,

$$U_{\text{ФЭМ}} = \frac{e a r D B (\mu_n + \mu_p) \Delta n(0)}{l \sigma_0 + e (\mu_n + \mu_p) \Delta n(0) L}. \quad (13.54)$$

При высоком уровне возбуждения, когда $\Delta n \gg (n_0 + p_0)$, в (13.54) можно пренебречь членом $l \sigma_0$, а поэтому

$$U_{\text{ФЭМ}} = r a D B / L. \quad (13.55)$$

При низком уровне возбуждения, когда $\Delta n \ll (n_0 + p_0)$, в знаменателе выражения (13.54) можно пренебречь вторым слагаемым. В результате получим:

$$U_{\text{ФЭМ}} = e r a D B (\mu_n + \mu_p) \frac{\Delta n(0)}{l \sigma_0}. \quad (13.56)$$

Если скорость поверхностной рекомбинации пренебрежимо мала, то полное число неравновесных носителей заряда в образце, входящих на единицу освещенной поверхности, определяется квантовым выходом β , интенсивностью возбуждающего света I и временем жизни носителей заряда τ . Следовательно, учитывая (13.53), можно написать:

$$\int_0^l \Delta n dx = \beta \tau \frac{I}{h\nu} = \Delta n(0) L, \quad (13.57)$$

откуда

$$\Delta n(0) = \frac{\beta \tau I}{h\nu L} = \frac{\beta L I}{h\nu D}. \quad (13.58)$$

Подставляя (13.58) в (13.56), получаем:

$$U_{\text{ФЭМ}} = e r a B (\mu_n + \mu_p) \beta L \frac{I}{h\nu l \sigma_0}. \quad (13.59)$$

Из анализа (13.59) и (13.55) следует, что напряжение ФЭМ-эффекта при малом уровне возбуждения пропорционально интенсивности света, а с ростом уровня освещения стремится к насыщению (рис. 13.14).

13.7. ФОТОЭФФЕКТ В p - n ПЕРЕХОДЕ

В p - n переходе, как известно (см. § 9.8), существует потенциальный барьер, обусловленный электрическим полем, которое появляется в результате диффузии основных носителей заряда через p - n переход. При термодинамическом равновесии положение уровня Ферми во всей системе постоянно и энергетическая схема p - n перехода имеет вид, изображенный на рис. 13.15, a . В этом случае токи обусловлены свободными носителями заряда, генерируемыми за счет теплового возбуждения, и в равновесии суммарный ток равен нулю. При принятом обозначении токов, как это сделано на рис. 13.15, a , условие равновесия запишется в виде

$$-J_n^{(n)} + J_n^{(p)} - J_p^{(p)} + J_p^{(n)} = 0. \quad (13.60)$$

В этом равенстве каждая пара токов электронов и дырок равна нулю

$$-J_n^{(n)} + J_n^{(p)} = 0; \quad -J_p^{(p)} + J_p^{(n)} = 0, \quad (13.61)$$

так как количества переходящих носителей заряда в прямом и обратном направлении при термодинамическом равновесии равны. Но потоки неосновных носителей заряда — электронов из p -области и дырок из n -области есть не что иное, как электронная и дырочная составляющие тока насыщения в вольт-амперной характеристике диода (9.154). Полный ток насыщения

$$J_s = J_{ns} + J_{ps} = \frac{eD_n n_p}{L_n} + \frac{eD_p p_n}{L_p}. \quad (13.62)$$

Рассмотрим теперь p - n переход, на который падают фотоны с энергией, большей, чем ширина запрещенной зоны (рис. 13.15, b). В ре-

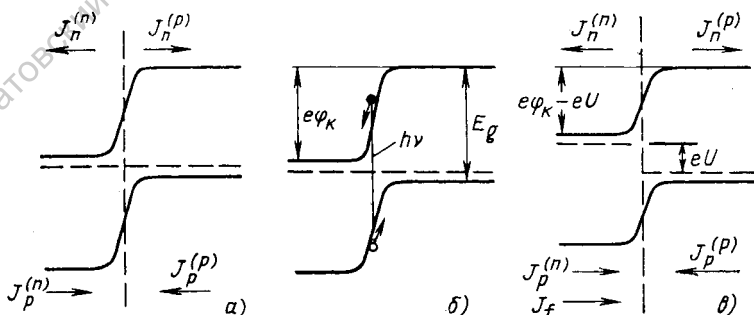


Рис. 13.15. Энергетическая схема p - n перехода и токи при термодинамическом равновесии (a) и при освещении (b , c)

в результате поглощения фотона возникает электронно-дырочная пара. Под действием внутреннего поля в p - n переходе созданные светом носители заряда движутся в противоположных направлениях: дырки — в p -область, а электроны — в n -область (см. рис. 13.15, б). Эти перешедшие через p - n переход неравновесные носители заряда создадут добавочную плотность тока, которую обозначим J_f . Так как перешедшие в p -область избыточные дырки уменьшают отрицательный объемный заряд, то энергетические уровни в p -области снижаются и в результате этого происходит понижение потенциального барьера. Следовательно, разделение зарядов приводит к возникновению разности потенциалов U (рис. 13.15, в). Электроны из n -области и дырки из p -области, преодолевая пониженный потенциальный барьер $e(\varphi_n - U)$, будут переходить соответственно в p - и n -области. При этом токи, обусловленные инжектированными носителями заряда, направлены из p - в n -область (рис. 13.15, в). Стационарное состояние установится тогда, когда число создаваемых светом электронно-дырочных пар сравняется с числом носителей заряда, уходящих через пониженный потенциальный барьер p - n перехода. Возникшая в p - n переходе фото-ЭДС носит название **вентильной**. Следовательно, освещенный p - n переход действует как фотозащелка.

Для определения вентильной фото-ЭДС $U_{\text{вен}}$ запишем уравнение для тока J , текущего через p - n переход:

$$J = J_f - J_n^{(n)} + J_n^{(p)} - J_p^{(p)} + J_p^{(n)}. \quad (13.63)$$

Здесь плотности токов неосновных носителей заряда при освещении $J_n^{(p)}$ и $J_p^{(n)}$, как следует из рис. 13.15, в, равны своим значениям в равновесии:

$$J_n^{(p)} = J_{sn}; \quad J_p^{(n)} = J_{sp}. \quad (13.64)$$

Плотности токов основных носителей заряда при освещении $J_n^{(n)}$ и $J_p^{(p)}$ в результате понижения потенциального барьера на eU увеличиваются и становятся равными:

$$J_n^{(n)} = J_{sn} e^{eU/kT}; \quad (13.65)$$

$$J_p^{(p)} = J_{sp} e^{eU/kT}. \quad (13.66)$$

Объединяя выражения (13.64)–(13.66), получаем:

$$J_f - J_{sn}(e^{eU/kT} - 1) - J_{sp}(e^{eU/kT} - 1) = J \quad (13.67)$$

или

$$J_f - J = J_s(e^{eU/kT} - 1), \quad (13.68)$$

откуда будем иметь:

$$U = \frac{kT}{e} \ln \left(1 + \frac{J_f - J}{J_s} \right). \quad (13.69)$$

Уравнение (13.69) является уравнением фотодиода для любого режима.

Для определения вентильной фото-ЭДС $U_{\text{вен}}$, которая соответствует напряжению на зажимах разомкнутой цепи, необходимо положить $J = 0$. Тогда

$$U_{\text{вен}} = \frac{kT}{e} \ln \left(1 + \frac{J_f}{J_s} \right). \quad (13.70)$$

Значение J_f определяется числом избыточных носителей заряда, созданных светом и дошедших до p - n перехода. Если через $I/h\nu$ обозначим число фотонов, падающих каждую секунду на единицу поверхности, через β — квантовый выход, т. е. количество электронно-дырочных пар, возникающих на один фотон, а через γ — долю непрорекомбинировавших пар носителей заряда, пришедших к p - n переходу, то

$$J_f = e\gamma\beta \frac{I}{h\nu}. \quad (13.71)$$

Это выражение справедливо для случая, когда весь падающий на полупроводник свет поглощается. С учетом (13.71) выражение (13.70) примет вид:

$$U_{\text{вен}} = \frac{kT}{e} \ln \left(1 + \frac{e\gamma\beta I}{h\nu J_s} \right). \quad (13.72)$$

При большом уровне освещения, когда $J_f/J_s \gg 1$, имеем:

$$U_{\text{вен}} = \frac{kT}{e} \ln \frac{e\gamma\beta I}{h\nu J_s}. \quad (13.73)$$

При малом уровне возбуждения, когда $J_f/J_s \ll 1$, используя разложение в ряд, получаем

$$U_{\text{вен}} = kT \frac{\gamma\beta I}{h\nu J_s}, \quad (13.74)$$

т. е. вентильная фото-ЭДС при малом уровне возбуждения пропорциональна интенсивности света.

Полученные зависимости вентильной фото-ЭДС от интенсивности возбуждающего света достаточно хорошо согласуются с экспериментальными данными, как это следует из рис. 13.16 и 13.17, на

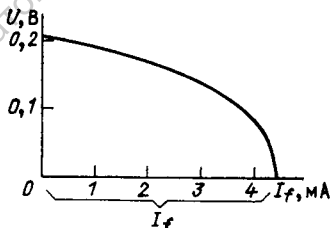


Рис. 13.16. Вольт-амперная характеристика в вентильном режиме фотодиода

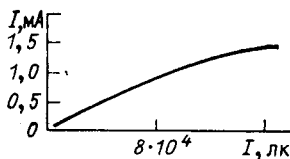


Рис. 13.17. Люкс-амперная характеристика в вентильном режиме германиевого фотодиода

которых приведены вольт-амперная и люкс-амперная характеристики для германиевого фотодиода в вентильном режиме. Из рис. 13.16 следует, что малым токам соответствует предельное для данной освещенности значение фото-ЭДС.

Из формулы (13.73) следует, что при увеличении интенсивности света фото-ЭДС возрастает до тех пор, пока не сравняются плотности токов J_f и J_s , т. е. пока не исчезнет потенциальный барьер, препятствующий переходу носителей заряда. Высота барьера представляет собой максимально достижимое значение $U_{\text{вен}}$. Поэтому вентильная фото-ЭДС зависит от степени легирования, а следовательно, от положения уровня Ферми. Практически этот предел соответствует ширине запрещенной зоны.

13.8. ФОТОЭФФЕКТ НА БАРЬЕРЕ ШОТТКИ

В контакте металла с полупроводником имеет место перераспределение зарядов, приводящее к возникновению в приповерхностной области полупроводника обедненного слоя, называемого барьером Шоттки. Если работа выхода из металла Φ_m больше работы выхода из полупроводника n -типа Φ_n (рис. 13.18, а), образуется барьер Шоттки для электронов (рис. 13.18, б). Если $\Phi_m < \Phi_n$, то образуется барьер Шоттки для дырок. Высота этого барьера равна:

$$e\varphi_K = \Phi_m - \Phi_n. \quad (13.75)$$

Предположим, что на структуру, энергетическая схема которой изображена на рис. 13.18, б, падают фотоны с энергией $h\nu > e\varphi_K$. В результате оптического возбуждения электроны в металле приобретают энергию, достаточную для преодоления барьера (процесс 1). Электроны, входя в полупроводник, заряжают его отрицательно, т. е. на барьере возникает фото-ЭДС. Электроны, возбужденные из валентной зоны полупроводника в металл, будут в металле занимать состояния выше уровня Ферми (процесс 2). Но этот переход совершается при помощи туннелирования, его вероятность намного меньше, чем для процесса 1. Кроме того, за счет процесса 2 не будет возникать фото-ЭДС, поскольку дырки, переходя в металл, будут рекомбинировать с электронами вблизи границы с металлом.

Если $h\nu > E_g$, то электронно-дырочные пары, генерируемые в пределах барьера Шоттки (процесс 3), разделяются контактным

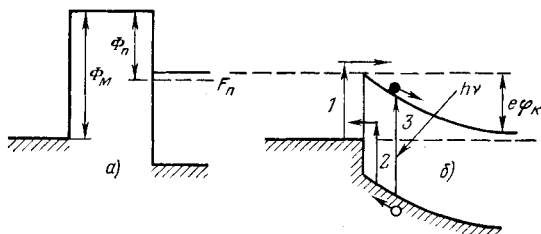


Рис. 13.18. Образование барьера Шоттки в полупроводнике, находящемся в контакте с металлом, и возможные процессы внутренней фотоэмиссии

полем и создают фото-ЭДС между металлическим электродом и объемом полупроводника.

На рис. 13.19 приведена спектральная зависимость фототока контакта металл—полупроводник. Видно, что на характеристике имеется две области: область большого фототока, обусловленного генерацией пар носителей заряда при межзонных переходах, и область меньших значений фототока, простирающаяся в длинноволновую сторону и связанная с миграцией горячих электронов из металла в полупроводник с переходом через барьер.

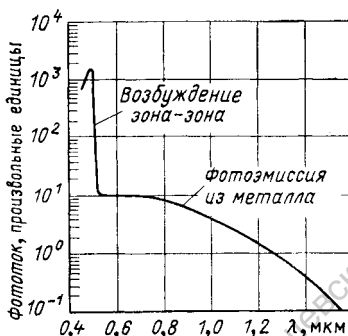


Рис. 13.19. Спектральная зависимость фототока контакта металл—полупроводник

Фотодиоды с барьером Шоттки могут быть выполнены в виде фотоэлементов большой площади. Они могут быть использованы для регистрации фототока, вызванного α -частицами. Кроме того, они могут применяться для регистрации электронов, возбужденных рентгеновскими лучами, γ -лучами и тепловыми нейтронами. Тепловые нейтроны представляют собой неионизирующие частицы, но, например, в InP они вызывают превращение In^{115} в радиоактивный изотоп In^{116} , который превращается в Sn^{116} в результате β -распада с эмиссией электронов высокой энергии, которые и генерируют электронно-дырочные пары.

13.9. ВНЕШНИЙ ФОТОЭФФЕКТ

На рис. 13.20 представлена энергетическая схема полупроводника разной степени легирования. Здесь $E_{\text{вак}}$ есть энергия, которой обладает электрон, вышедший из полупроводника и имеющий в вакууме практически нулевую кинетическую энергию. Энергия χ , отделяющая край зоны проводимости от уровня вакуума, есть энергия электронного сродства. Величина Φ , равная разности энергий,

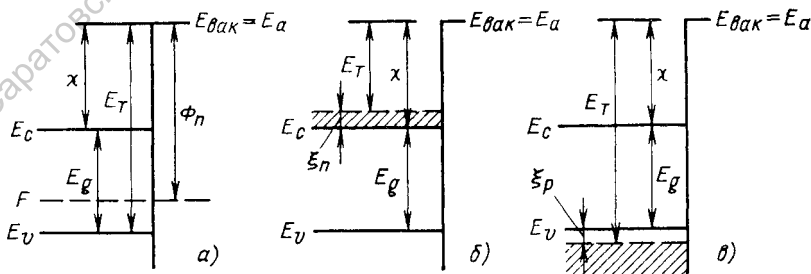


Рис. 13.20. Зависимость внешнего фотоэффекта от степени легирования полупроводника

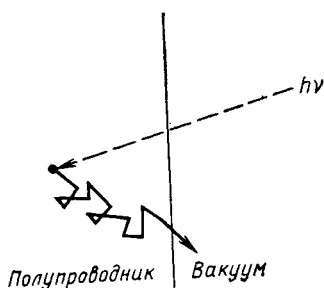


Рис. 13.21. Возбуждение, рассеяние и выход фотоэлектронов из полупроводника

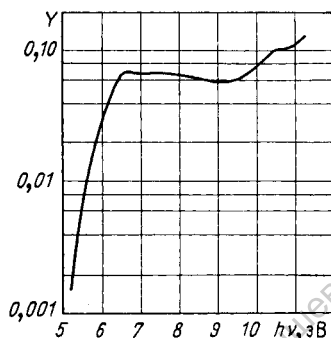


Рис. 13.22. Спектральное распределение квантового выхода электронов из CdTe. Порог фотоэффекта равен примерно 5 эВ

соответствующих уровню Ферми и уровню вакуума, есть работа выхода.

Рассмотрим теперь взаимодействие между фотоном и полупроводником, в результате которого происходит эмиссия электрона из полупроводника. Процесс эмиссии электронов из полупроводника под действием излучения называют *внешним фотоэффектом*. Внешний фотоэффект представляет собой последовательность трех процессов (рис. 13.21): 1) электрон валентной зоны полупроводника переходит в высокое энергетическое состояние зоны проводимости в результате взаимодействия с фотоном; 2) возбужденный электрон в результате рассеяния теряет часть энергии и переходит на нижний уровень зоны проводимости; 3) электрон выходит с нижнего уровня зоны проводимости полупроводника в вакуум с энергией, равной разности его полной энергии и $E_{\text{вак}}$. Порог внешнего фотоэффекта E_T есть наименьшая энергия фотона, которая достаточна, чтобы удалить электрон из полупроводника.

У собственного или невырожденного полупроводника, когда уровень Ферми находится в запрещенной зоне (рис. 13.20, а), эмиссия электронов под действием света происходит из валентной зоны. Поэтому

$$E_T = \chi + E_g. \quad (13.76)$$

Для непрямых переходов, когда сохранение квазинимпульса обеспечивается за счет эмиссии фонона с энергией E_p ,

$$E_T = \chi + E_g + E_p. \quad (13.77)$$

Для сильно легированного полупроводника *n*-типа, у которого уровень Ферми лежит выше края зоны проводимости на величину ξ_n (рис. 13.20, б), имеем:

$$E_T = \chi - \xi_n. \quad (13.78)$$

В сильно легированном полупроводнике p -типа уровень Ферми расположен на величину ξ_p ниже края валентной зоны (рис. 13.20, в), поэтому

$$E_T = \chi + E_g + \xi_p. \quad (13.79)$$

Работа выхода Φ , порог фотоэффекта E_T и порог для прямых переходов E_{Td} для некоторых полупроводников приведены в табл. 13.1. В ней даны также значения ширины запрещенной зоны E_g , энергия электронного сродства χ и указаны плоскости, с которых происходил фотоэффект.

Т а б л и ц а 13.1. Работа выхода и порог внешнего фотоэффекта для некоторых полупроводников

Полупроводник и плоскость	E_g , эВ	Φ , эВ	E_T , эВ	E_{Td} , эВ	$\chi = E_T - E_g$, эВ
AlSb (110)	1,5	4,86	5,15	5,75	3,65
GaAs (110)	1,40	4,71	5,47	5,75	4,07
GaSb (110)	0,70	4,76	4,76	5,24	4,06
InAs (110)	0,41	4,90	5,31	5,58	4,90
InSb (110)	0,18	4,75	4,77	5,26	4,59
InP (110)	1,3	4,45	5,68	5,94	4,38
Si (111)	1,09	4,83	5,10	5,45	4,01
Ge (111)	0,67	4,80	4,80	5,22	4,13

Типичная кривая зависимости квантового выхода фотоэлектронов от энергии фотонов приведена на рис. 13.22. Вначале по мере увеличения энергии фотонов, превышающей пороговое значение, число эмиттируемых электронов возрастает и квантовый выход резко нарастает. Затем он переходит в плато, на котором имеется структура, отражающая свойства зонной структуры полупроводника. Начальный участок резкого нарастания кривой квантового выхода $Y(h\nu)$ описывается зависимостью вида

$$Y(h\nu) = A(h\nu - E_T)^r, \quad (13.80)$$

где A и r — константы, r определяется механизмом рассеяния.

В случае прямого перехода электрона из валентной зоны без рассеяния $r = 1$, а при наличии рассеяния $r = 2$. Для непрямого перехода электрона из валентной зоны без рассеяния и с рассеянием $r = 5/2$.

ПРИЛОЖЕНИЯ

1. Свойства Ge, Si и GaAs (при 300 К)

Свойства	Ge	Si	GaAs
Количество атомов в 1 см ³	4,42·10 ²²	5,0·10 ²²	2,21·10 ²²
Атомная масса	72,6	28,08	144,63
Поле пробоя, В/см	~10 ⁵	~3·10 ⁵	~4·10 ⁵
Кристаллическая структура	Алмаза	Алмаза	Цинковой обманки
Плотность, г/см ³	5,3267	2,328	5,32
Относительная диэлектрическая проницаемость	16	11,8	10,9
Эффективная плотность состояний в зоне проводимости N_c , см ⁻³	1,04·10 ¹⁹	2,8·10 ¹⁹	4,7·10 ¹⁷
Эффективная плотность состояний в валентной зоне N_v , см ⁻³	6,1·10 ¹⁸	1,02·10 ¹⁹	7,0·10 ¹⁸
Эффективные массы m^*/m_0 : электронов	m_e 1,59 m_i 0,082	m_e 0,92 m_i 0,19	0,068
дырок	$m_{рл}^*$ 0,04 $m_{рт}^*$ 0,30	$m_{рл}^*$ 0,16 $m_{рт}^*$ 0,50	$m_{рл}^*$ 0,12 $m_{рт}^*$ 0,45
Сродство к электрону, эВ	4,0	4,05	4,07
Ширина запрещенной зоны при 300 К, эВ	0,66	1,11	1,43
Собственная концентрация, см ⁻³	2,5·10 ¹³	1,6·10 ¹⁰	1,1·10 ⁷
Постоянная решетки, нм	0,565748	0,543086	0,56534
Температурный коэффициент линейного расширения $\Delta L/L \Delta T$, °C ⁻¹	5,8·10 ⁻⁶	2,6·10 ⁻⁶	5,9·10 ⁻⁶
Точка плавления, °C	937	1420	1238
Время жизни неосновных носителей, с	10 ⁻⁸	2,5·10 ⁻⁸	10 ⁻⁸
Подвижность дрейфовая, см ² /(В·с): электронов μ_n дырок μ_p	3900 1900	1500 600	8500 400
Энергия рамановских фононов E_p , эВ	0,037	0,063	0,035
Удельная теплота, Дж/(г·°C)	0,31	0,7	0,35
Теплопроводность (при 300 К), Вт/(см·°C)	0,64	1,45	0,46
Коэффициент тепловой диффузии, см ² /с	0,36	0,9	0,44
Давление паров, Па	0,133 при 1270 °C 1,33·10 ⁻⁶ при 800 °C	0,133 при 1600 °C 1,33·10 ⁻⁶ при 930 °C	1,01·10 ⁶ при 1050 °C 1,01·10 ⁻⁷ при 1220 °C
Работа выхода, эВ	4,4	4,8	4,7

II. Свойства полупроводников

Полупроводники	Ширина запрещенной зоны		Тип абсолютного минимума зоны проводимости	$\left(\frac{dE}{dT}\right)_{(300\text{ K})} \cdot 10^4$, эВ/К	$\left(\frac{dE}{dP}\right)_{(300\text{ K})} \cdot 10^4$, эВ/Па	Эффективная масса		Показатель преломления n	Относительная диэлектрическая проницаемость ϵ_r	Постоянная ре-шетки a , нм	Подвижность	
	E_g (0 К), эВ	E_g (300 К), эВ				m_n^*/m_0	m_p^*/m_0				μ_n , см ² /В·с	μ_p , см ² /В·с
Si	1,166	1,12	Непрямой 100	-2,3	-1,5	$m_n^* 0,92$	$m_p^* 0,56$	3,44	11,7	0,543	1350	480
Ge	0,74	0,67	Непрямой 111	-3,7	5,0	$m_n^* 0,19$	$m_p^* 0,16$	4,00	16,3	0,566	3900	1900
α -Sn	-0,2		Прямой 000		5,0	$m_n^* 1,59$	$m_p^* 0,33$			0,6489	2000	1000
α	0 (6H)	2,8-3,2	Непрямой	-3,3		$m_n^* 0,082$	$m_p^* 0,04$	2,69 с		а0,30817 с1,51123	400	
β	2,68	2,2	Непрямой			0,02		2,65 $\perp$ с	10,2	0,4359		
Se	1,95	1,74	Прямой 0001	-14	-20		0,12	5,56 с 3,72 $\perp$ с	8,5		1	
Te	0,334	0,32	Прямой 0001	-0,3	-19	0,038 $\perp$	0,26 0,10 $\perp$	3,07 с 2,68 $\perp$ с	5,0 с 2,2 $\perp$ с		1100	
BP		2	Непрямой					2,6	6,9	0,4538	80	
AlP	2,5	2,43	Непрямой 100	-3,5	0,13			3,0	9,8	0,5462		
AlAs	2,24	2,16	Непрямой		0,5	$m_n^* 1,06$			12	0,566	1000	~100
AlSb	1,6	1,6	Непрямой 100	-4	-1,6	$m_n^* 0,49$	0,39	3,4	11	0,6135	50	400

Продолжение

Полупроводники	Ширина запрещенной зоны		Тип абсолютного минимума зоны проводимости	$\left(\frac{dE}{dT}\right)_{300\text{ K}}$, эВ/К	$\left(\frac{dE}{dT}\right)_{300\text{ K}}$, эВ/К	Эффективная масса		Показатель преломления n	Относительная диэлектрическая проницаемость ϵ	Постоянная решетка a , нм	Подвижность	
	E_g (0 К), эВ	E_g (300 К), эВ				m_n^*/m_0	m_p^*/m_0				μ_n , см ² /(В·с)	μ_p , см ² /(В·с)
III—V	GaN	3,6	3,5	Прямой 0000	—3,9	3,7	0,2	2,4		a0,318 с0,516	150	120
	GaP	2,4	2,25	Непрямой 100	—5,4	—1,7	0,13	3,37	10	0,5450	120	400
	GaAs	1,520	1,43	Прямой 000	—5,0	11	0,07	3,4	12	0,5653	8600	650
	GaSb	0,81	0,69	Прямой 000	—4,1	12	0,045	3,9	15	0,6095	4000	650
	InP	1,42	1,28	Прямой 000	—4,6	4,6	0,07	3,37	12,1	0,58687	4000	240
	InAs	0,43	0,36	Прямой 000	—3,3	5	0,028	3,42	12,5	0,6058	30000	5000 (78 К)
	InSb	0,235	0,18	Прямой 000	—2,9	15	0,0133	3,75	18	0,64787	76000	180
	ZnO	3,43	3,3	Прямой 0000	—9,5	0,6	0,32	2,2	8,5	a0,32496 с0,52065	200	28
II—VI	ZnS α	3,91	3,78	Прямой 0000	—3,8	9	0,25	2,4	8,3	a0,5409		
	ZnS β	3,84	3,71	Прямой 000	—5,3	5,7	0,39	2,4	8,3	a0,3826 с0,6261		
	ZnSe	2,81	2,67	Прямой 000	—7,2	6	0,17	2,89	8,1	0,5667	530	

Полупроводники	Ширина запрещенной зоны		Тип абсолютного минимума зоны проводимости	$\left(\frac{dE_g}{dK}\right)_{(300\text{ K})} \cdot 10^4$, эВ/К	$\left(\frac{dE_g}{dK}\right)_{(300\text{ K})} \cdot 10^4$, эВ/К	Эффективная масса	Показатель преломления n_0	Относительная электронная проводимость μ_n	Постоянная ре-сетки a_0 , нм	Подвижность	
	E_g (0 К), эВ	E_g (300 К), эВ				m_{Hl}^*/m_0				μ_n , см ² /В·с	μ_p , см ² /В·с
ZnTe	2,39	2,28	Прямой 000	-5	6	0,15	3,56	9,7	0,6101	350	15
CdS	2,58	2,53	Прямой 0000	-5	3,3	0,20	2,5 0,7 $\perp$ c 5 $\parallel$ c	8,96	0,4136 0,6713	500	
CdSe	1,85	1,74	Прямой 0000	-4,6		0,13	2,5 $\parallel$ 0,4 $\perp$	10,6	0,4299 0,7010	600	
CdTe	1,60	1,50	Прямой 000	-4,1	1,5	0,11	0,35	10,9	0,6477		
HgS	2,5									5500	
HgSe	-0,24	-0,15	Прямой 000			0,045		25	0,6085		
HgTe	-0,28	-0,15	Прямой 000	+5,6		0,029	~0,3	20	0,642	22 000	100 (20 K)
PbS	0,29	0,41	Прямой 111	+4	-7	m_l 0,105 m_t 0,08	3,7	170	0,5936	550	600
PbSe	0,165	0,27	Прямой 111	+4	-8	m_l 0,07 m_t 0,039	3,7	250	0,6124	1020	930
PbTe	0,19	0,31	Прямой 111	+4	-9	m_l 0,24 m_t 0,024	3,8	412	0,6460	1620	750
SnTe	0,3	0,18	Прямой 111	-	-	-	-		0,6328	-	-

II-VI

IV-VI

III. Физические константы

Заряд электрона $e = 4,8 \cdot 10^{-10}$ ед. CGSE = $1,602 \cdot 10^{-19}$ Кл

Масса свободного электрона $m_0 = 9,11 \cdot 10^{-28}$ г

Скорость света в вакууме $c = 2,998 \cdot 10^{10}$ см/с

Боровский радиус $a_0 = 5,29 \cdot 10^{-9}$ см

Постоянная Планка $h = 6,62 \cdot 10^{-34}$ Дж·с = $4,5 \cdot 10^{-15}$ эВ·с

Постоянная Дирака $\hbar = h/2\pi = 1,054 \cdot 10^{-34}$ Дж·с = $7,1 \cdot 10^{-16}$ эВ·с

Постоянная Больцмана $k = 1,380 \cdot 10^{-23}$ Дж/К = $8,62 \cdot 10^{-5}$ эВ/К

Тепловая энергия kT

25,9 мэВ при комнатной температуре

6,7 мэВ при температуре жидкого азота

0,36 мэВ при температуре жидкого гелия

Энергия, соответствующая 1 эВ, 1 эВ = $1,602 \cdot 10^{-19}$ Дж

Длина волны в вакууме излучения с энергией фотона 1 эВ λ_0 (1 эВ) = $1,239 \times 10^{-4}$ см

Волновое число, соответствующее 1 эВ, ν_0 (1 эВ) = $8,06 \cdot 10^4$ см⁻¹

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

А

- Адиабатическое приближение 24
- Аккумуляция носителей заряда 236
- Акустические ветви колебаний решетки 78
- Акцептор, определение 18
- Акцепторный полупроводник 18, 119
- Амбиполярная диффузионная подвижность 239
- дрейфовая подвижность 238
- Ангармонизм колебаний 91
- Ангармонический осциллятор 92
- Антизапорный слой 250

Б

- Барьер Шоттки 285
- Барьерная емкость 264
- Бимолекулярное рекомбинационное свечение 345
- Биолуминесценция 336
- Биполярная оптическая генерация носителей заряда 202

В

- Вероятность переходов 135
- поглощения фонона 209, 306
- рассеяния 135
- релаксации 147
- Вертикальные переходы 305
- Виды рекомбинации 336
- Влияние давления 317
- поля магнитного 318
- электрического 318
- температуры 316
- Внешний фотоэффект 375
- Внешняя контактная разность потенциалов 248
- Внутренний фотоэффект 357
- Внутренняя контактная разность потенциалов 247
- Водородоподобные центры 66
- Волновое число 28, 71
- Волновой вектор пакета 48
- — решетки 72

- Волновой вектор фонона 83
- — электрона 28
- Вольт-амперная характеристика контакта металл—полупроводник 255, 260
- — — *p-n* перехода 266
- Время жизни дырки 208, 217
- — мгновенное 204, 212
- — неравновесных носителей заряда 203, 208, 212, 218
- — температурная зависимость 219
- — фотона 210, 306
- — электрона 208, 212, 217
- — электронно-дырочной пары 203, 212, 217
- релаксации 139, 152, 153
- максвелловское 205
- при рассеянии на ионах примеси 160
- — — решеточном рассеянии 160
- Выпрямление на контакте металл — полупроводник 253
- — *p-n* переходе 264
- Вырожденный примесный полупроводник 106
- собственный полупроводник 112

Г

- Гармонический осциллятор 76, 81
- Генерация носителей заряда 13, 199
- — — биполярная 202
- — — монополярная 204
- Гетеропереход 275

Д

- Демаркационный уровень 223
- Дефекты 142
- линейные 142
- точечные 142
- Диодная теория выпрямления тока 256
- Дислокации 142
- Диффузионная длина 234
- скорость 234
- теория выпрямления тока 258
- Диффузионное рассеяние 292
- Диффузионный ток 226
- Диффузия носителей заряда 224, 229, 232

Диэлектрическое время релаксации 205
 Длина диффузионная 234
 — дрейфа 235
 — затягивания 234
 — свободного пробега носителей заряда 14, 142, 147, 152, 153
 — — — фонона 210, 306
 — экранирования 231, 243
 Долины 60
 Домен 192
 Донорно-акцепторные пары 344
 Донорный полупроводник 19, 114
 Доноры, определение 19
 Дрейфовая скорость носителей заряда 15, 21, 48
 Дрейфовый ток 226
 Дырки 13, 17
 — легкие 62, 63
 — тяжелые 62, 63

Е

Емкость контакта металл—полупроводник 252
 — *p-n* перехода 264

З

Закон Ома 186
 — сохранения квазимпульса 304
 — — энергии 304
 Запорный слой 250
 Зона Бриллюэна, первая 39
 — валентная 16
 — запрещенная 16
 — примесной проводимости 124
 — проводимости 16
 Зонная структура антимонида индия 64
 — — арсенида галлия 60
 — — германия 60
 — — кремния 60

И

Избыточная концентрация носителей заряда 201
 Изгиб зон 241
 Изоэнергетические поверхности 54
 — — сферические 55
 — — эллипсоидальные 54, 55
 Импульс фотона 209
 — электрона 50
 Инверсная заселенность 352
 Инверсный слой 242
 Индукция магнитная 164
 Инжекция 236, 265
 Интеграл столкновения 137
 Ионизация примесей 116

К

Католюминесценция 336
 Квазимпульс 37

Квазиуровень Ферми 201, 253
 Квантовые генераторы 353
 Квантовый выход излучения 346
 — — фотоионизации 361
 Кинетическая энергия решетки 75
 Кинетическое уравнение Больцмана 133
 Ковалентные кристаллы 12
 Колебания атомов решетки 69, 70, 76
 — струны 69
 Компоненты тензора 52
 Контакт вырожденных электронного и дырочного полупроводников 277
 — металл—металл 246
 — металл—полупроводник 248
 — электронного и дырочного полупроводников 260
 Контактная разность потенциалов 179, 247, 248, 249
 Концентрация дырок 102, 104, 107
 — носителей заряда 93, 101
 — — — вырождения 108
 — — — зависимость от температуры 111, 118
 — электронов 101, 104, 107
 Коэффициент амбиполярной диффузии 238
 — диффузии 227
 — захвата 214
 — ионизации 214
 — отражения 302
 — Пельтье 181
 — поглощения 210
 — пропускания 303
 — рекомбинации 200
 — теплового расширения 91
 — теплопроводности 183
 — Томсона 181
 — Холла 166, 170
 — экстинкции 328

Л

Лавинный пробой 270
 Лазеры 353
 Ловушки захвата 213, 222
 — рекомбинации 213, 222
 Люминесценция 336
 — гашение 346
 — мономолекулярная 337
 — рекомбинационная 337

М

Магнитная проницаемость 328
 Максвелловское время релаксации 204
 МДП-структура 293

Н

Наклон зон 46, 107
 Невырожденный примесный полупроводник 8, 104

Невырожденный собственный полупроводник 109
 Некристаллические полупроводники 127
 Непрямые переходы 309
 Неравновесная функция распределения 133, 154
 Неравновесные носители заряда 200
 n^+n переход 271
 Нормальные координаты решетки 74

О

Область ионизации примеси 117
 — — сильной 117
 — — слабой 116
 Обменный интеграл 32
 Образование хвостов плотности состояния 126
 Обращенный слой 242
 Одноэлектронное приближение 25
 Омический контакт 281
 Оператор Гамильтона 23
 Оптические ветви колебаний решетки 77

П

Переходы вертикальные 305
 — — внутризонные 332
 — — межзонные 304
 — — не прямые 309
 — — прямые 304
 Периодический потенциал решетки 31
 Плотность состояний 92
 — — тока 20
 — — дырочного 157
 — — электронного 157
 p - n переход 260
 — — физический 250
 p^+n переход 271
 Поверхностная проводимость 290
 — — рекомбинация 297
 Поверхностные состояния 296
 — — быстрые 296
 — — медленные 296
 — — уровни 282
 — — явления 282
 Поверхностный потенциал 286
 Поглощение примесное 304, 333
 — — решеточное 304, 334
 — — света 303
 — — свободными носителями заряда 327
 — — собственное 304, 309
 — — при не прямых переходах 309
 — — — — прямых переходах 304
 Подвижность носителей заряда 21, 159, 160
 — — при эффекте поля 292
 — — Холла 171
 Показатель поглощения 328
 — — преломления 328
 — — комплексный 328

Поле Холла 166
 Полупроводник 8
 — — акцепторный 19
 — — вырожденный 106, 112
 — — донорный 19
 — — компенсированный 120
 — — — — частично 120
 — — невырожденный 8, 104
 — — примесный 103
 — — собственный 109
 — — — — вырожденный 112
 — — невырожденный 109
 Поляризуемость 330
 Постоянная Больцмана 96
 — — Планка 23
 Потенциальная энергия решетки 75
 Правило отбора 305
 Приведенная масса 306
 Приведенный квазиуровень Ферми 201
 — — уровень Ферми 101
 Примесные зоны 126
 Принцип детального равновесия 137
 — — макроскопической обратимости 137
 — — Паули 37
 Проводимость 7, 157
 Процессы в p - n переходе при обратном смещении 265
 — — — — — прямом смещении 264
 — — — — — генерации 225
 — — переноса 134, 141
 — — рассеяния 137

Р

Работа выхода 244, 245, 246
 — — из акцепторного полупроводника 246
 — — — — — собственного полупроводника 246
 — — — — — электронного полупроводника 246
 Равновесная концентрация носителей заряда 107
 Равновесное состояние 138
 Равновесные носители заряда 9, 199
 Радиолюминесценция 336
 Разогрев электронно-дырочного газа 186
 Рассеяние диффузное 292
 — — междолинное 190
 — — на акустических фоновых 151
 — — атомах примеси 147
 — — дислокациях 147
 — — ионах примеси 143
 — — оптических фоновых 153
 — — тепловых колебаниях решетки 148
 — — типы 132
 — — угол 144

Рекомбинация безызлучательная 206
 — донорно-акцепторных пар 344
 — излучательная 206
 — межзонная 211
 — Оже 206
 — поверхностная 297
 — при переходе зона—примесь 342
 — ударная 211
 — фононная 206
 — фотонная 206
 — через ловушки 213
 Релаксация люминесценции 345
 — фотопроводимости 362

С

Скорость генерации 225
 — групповая 270
 — звуковая 270
 — поверхностной рекомбинации 297
 — рекомбинации 225
 — фазовая 270
 — фононная 270
 — фотонная 306
 Слой объемного заряда p - n перехода 263
 Собственная концентрация 110
 Соотношение Эйнштейна 228
 Соударения неупругие 141
 — упругие 141
 Спектр излучения 337
 — отражения 302
 — поглощения 303
 Спонтанное излучение 347
 Статистика Бозе—Эйнштейна 83
 — Больцмана 98
 — Ферми—Дирака 96
 — фононов 82
 Степень вырождения 100
 Стимулированное излучение 349, 352
 Сферические поверхности равной энергии 55

Т

Температура вырождения 108
 — Дебая 87, 88, 89
 — насыщения 117
 — появления собственной проводимости 117
 Теория выпрямления тока 253
 — — — диодная 256
 — — — диффузионная 258
 Тепловое расширение 90
 — сопротивление 90
 Теплоемкость 84
 Теплопроводность 183
 Ток насыщения 255, 258, 269
 Толщина объемного заряда 252, 255
 Триболюминесценция 336
 Туннельный диод 277
 — эффект 257

У

Угол Холла 167
 Ударная ионизация 186, 194
 — рекомбинация 211
 Уровень Ферми 113, 248
 — — зависимость от температуры 113
 Уровни глубокое 69
 — Ландау 321
 — Тамма 282
 Условие цикличности Борна—Кармана 35

Ф

Фононы 82
 — акустические 84
 — оптические 84
 Фотолюминесценция 336
 Фотопроводимость 360
 Фотоэлектромагнитный эффект 368
 Фотоэффект 371
 — внешний 375
 — внутренний 357
 Функция Блоха 29
 — Больцмана 98
 — Ферми—Дирака 96

Х

Хвосты зон 126
 Хемилюминесценция 336
 Холл-фактор 170

Ц

Циклотронная частота 58
 Циклотронный резонанс 57

Ч

Число состояний 35

Ш

Ширина запрещенной зоны 16, 112, 306
 — — — зависимость от давления 317
 — — — — температуры 316
 ЭДС Дембера 367
 — термоэлектродвижущая 177
 Экситонное излучение 340
 — поглощение 323
 Экситонные комплексы 326
 Экситоны 323
 — не прямые 326
 — прямые 326
 — свободные 325
 — связанные 326
 Эксклюзия носителей заряда 236
 Экстракция носителей заряда 236
 Электролюминесценция 336

Электропроводность примесного полупроводника 18
— собственного полупроводника 12
Электростатическая ионизация 186, 197
Элементы тензора 52
Эллипсоидальные поверхности равной энергии 54, 93
Энергетическая структура *p-n* перехода 261
— щель 16
Энергия активации 106, 111
— гармонического осциллятора 76
— ионизации примеси 67
— связи экситона 324
— Ферми 96
— фонона 83
— электронного сродства 244
Эффект Ганна 186, 190
— Дембера 370

Эффект Зеебека 177
— магнетопоглощения 322
— магниторезистивный 172
— Пельтье 177
— поля 290
— Томсона 177
— фононного увлечения 180
— фотоэлектромагнитный 368
— Франца—Келдыша 318
— Холла 164, 167
Эффективная масса 51
— — дырки 62
— — — легкой 63
— — — тяжелой 63
— — плотности состояний 101, 102
— — поперечная 61
— — продольная 61
Эффективное сечение захвата 217
— — проводимости 145
— — рассеяния 131

ОГЛАВЛЕНИЕ

Предисловие	3
Список основных обозначений	4
Глава первая. Полупроводники. Элементарная теория электропроводности	7
1.1. Классификация веществ по удельной электрической проводимости Полупроводники	7
1.2. Модельные представления о механизме электропроводности собственных полупроводников	12
1.3. Модельные представления о механизме электропроводности примесных полупроводников	18
1.4. Элементарная теория электропроводности полупроводников	20
Глава вторая. Основы зонной теории полупроводников	22
2.1. Уравнение Шредингера для кристалла	22
2.2. Адиабатическое приближение и валентная аппроксимация	24
2.3. Одноэлектронное приближение	25
2.4. Приближение сильно связанных электронов	29
2.5. Число состояний электронов в энергетической зоне	35
2.6. Квазимпульс	37
2.7. Зоны Бриллюэна	38
2.8. Возможное заполнение электронных состояний валентной зоны	40
2.9. Зависимость энергии электрона от волнового вектора у дна и потолка энергетической зоны	42
2.10. Движение электронов в кристалле под действием внешнего электрического поля	45
2.11. Эффективная масса носителей заряда	51
2.12. Циклотронный резонанс	57
2.13. Зонная структура некоторых полупроводников	59
2.14. Метод эффективной массы	64
2.15. Элементарная теория примесных состояний	66
Глава третья. Колебания атомов кристаллической решетки	69
3.1. Одномерные колебания однородной струны	69
3.2. Колебания одноатомной линейной цепочки	70
3.3. Энергия колебаний атомов одномерной решетки. Нормальные координаты	74
3.4. Колебания двухатомной линейной цепочки	76
3.5. Колебания атомов трехмерной решетки	79
3.6. Статистика фононов	82
3.7. Теплоемкость кристаллической решетки	84
3.8. Термическое расширение и тепловое сопротивление твердого тела	90

Глава четвертая. Статистика электронов и дырок в полупроводниках	92
4.1. Плотность квантовых состояний	92
4.2. Функция распределения Ферми—Дирака	96
4.3. Степень заполнения примесных уровней	98
4.4. Концентрации электронов и дырок в зонах	100
4.5. Примесный полупроводник	103
4.6. Собственный полупроводник	109
4.7. Зависимость уровня Ферми от концентрации примеси и температуры для невырожденного полупроводника	113
4.8. Зависимость уровня Ферми от температуры для невырожденного полупроводника с частично компенсированной примесью	120
4.9. Примесные полупроводники при очень низких температурах	124
4.10. Некристаллические полупроводники	127
Глава пятая. Рассеяние электронов и дырок в полупроводниках	131
5.1. Механизмы рассеяния электронов и дырок	131
5.2. Кинетическое уравнение Больцмана	133
5.3. Равновесное состояние	139
5.4. Время релаксации	140
5.5. Рассеяние на ионах примеси	143
5.6. Рассеяние на атомах примеси и дислокациях	147
5.7. Рассеяние на тепловых колебаниях решетке	148
Глава шестая. Кинетические явления в полупроводниках	154
6.1. Неравновесная функция распределения	154
6.2. Удельная электрическая проводимость полупроводников	157
6.3. Зависимость подвижности носителей заряда от температуры	160
6.4. Эффект Холла	
6.5. Эффект Холла в полупроводниках с двумя типами носителей заряда	167
6.6. Магниторезистивный эффект	172
6.7. Термоэлектрические явления	177
6.8. Теплопроводность полупроводников	183
6.9. Электропроводность полупроводников в сильном электрическом поле	185
6.10. Эффект Ганна	190
6.11. Ударная ионизация	194
6.12. Туннельный эффект и электростатическая ионизация	197
Глава седьмая. Генерация и рекомбинация электронов и дырок	199
7.1. Равновесные и неравновесные носители заряда	199
7.2. Биполярная оптическая генерация носителей заряда	202
7.3. Монополярная оптическая генерация носителей заряда. Максвелловское время релаксации	204
7.4. Механизмы рекомбинации	205
7.5. Межзонная излучательная рекомбинация	206
7.6. Межзонная ударная рекомбинация	211
7.7. Рекомбинация носителей заряда через ловушки	213
7.8. Температурная зависимость времени жизни носителей заряда при рекомбинации через ловушки	219
7.9. Центры захвата и рекомбинационные ловушки	222
Глава восьмая. Диффузия и дрейф неравновесных носителей заряда	224
8.1. Уравнение непрерывности	224
8.2. Диффузионный и дрейфовый токи	226
8.3. Соотношение Эйнштейна	228

8.4. Диффузия и дрейф неравновесных носителей заряда в случае монополярной проводимости	229
8.5. Диффузия и дрейф неосновных избыточных носителей заряда в примесном полупроводнике	232
8.6. Диффузия и дрейф неравновесных носителей заряда в полупроводнике с проводимостью, близкой к собственной	236
Глава девятая. Контактные явления в полупроводниках	240
9.1. Полупроводник во внешнем электрическом поле	240
9.2. Термоэлектронная работа выхода	244
9.3. Контакт металл—металл. Контактная разность потенциалов	246
9.4. Контакт металл—полупроводник	248
9.5. Выпрямление тока в контакте металл—полупроводник	253
9.6. Диодная теория выпрямления тока	256
9.7. Диффузионная теория выпрямления тока	258
9.8. Контакт электронного и дырочного полупроводников	260
9.9. Выпрямление тока в p - n переходе	264
9.10. Теория тонкого p - n перехода	266
9.11. n^+ - n и p^+ - p переходы	271
9.12. Гетеропереходы	275
9.13. Контакт вырожденных электронного и дырочного полупроводников. Туннельный диод	277
9.14. Омический переход	281
Глава десятая. Поверхностные явления в полупроводниках	282
10.1. Природа поверхностных уровней	282
10.2. Теория слоя пространственного заряда	285
10.3. Эффект поля	290
10.4. Скорость поверхностной рекомбинации	297
10.5. Влияние поверхностной рекомбинации на время жизни носителей заряда в образцах конечных размеров	300
Глава одиннадцатая. Поглощение света полупроводниками	302
11.1. Спектр отражения и спектр поглощения	302
11.2. Собственное поглощение при прямых переходах	304
11.3. Собственное поглощение при непрямых переходах	309
11.4. Поглощение сильно легированного и аморфного полупроводников	313
11.5. Влияние внешних воздействий на собственное поглощение полупроводников	316
11.6. Экситонное поглощение	323
11.7. Поглощение свободными носителями заряда	327
11.8. Примесное поглощение	333
11.9. Решеточное поглощение	334
Глава двенадцатая. Люминесценция полупроводников	336
12.1. Типы люминесценции	336
12.2. Мономолекулярное свечение твердых тел	337
12.3. Рекомбинационное излучение полупроводников при фундаментальных переходах	337
12.4. Рекомбинационное излучение при переходах между зоной и примесными уровнями	341
12.5. Релаксация люминесценции полупроводников	345
12.6. Температурное тушение люминесценции полупроводников	346
12.7. Спонтанное и вынужденное излучение атома	347
12.8. Стимулированное излучение твердых тел	352

Глава тринадцатая. Фотоэлектрические явления в полупроводниках	357
13.1. Внутренний фотоэффект	357
13.2. Фотопроводимость	360
13.3. Релаксация фотопроводимости	362
13.4. Фотопроводимость при наличии поверхностной рекомбинации и диффузии носителей заряда	364
13.5. Эффект Дембера	366
13.6. Фотоэлектромагнитный эффект	368
13.7. Фотоэффект в p - n переходе	371
13.8. Фотоэффект на барьере Шоттки	374
13.9. Внешний фотоэффект	375
Приложения:	
I. Свойства Ge, Si и GaAs (при 300 K)	378
II. Свойства полупроводников	379
III. Физические константы	382
Предметный указатель	383

Клавдия Васильевна ШАЛИМОВА

**ФИЗИКА
ПОЛУПРОВОДНИКОВ**

Учебник

Издание четвертое,
стереотипное

ЛР № 065466 от 21.10.97

Гигиенический сертификат 78.01.07.953.П.007216.04.10
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»

lan@lanbook.ru; www.lanbook.com

192029, Санкт-Петербург, Общественный пер., 5.

Тел./факс: (812)412-29-35, 412-05-97, 412-92-72.

Бесплатный звонок по России: 8-800-700-40-71

Подписано в печать 30.07.10.

Бумага офсетная. Гарнитура Литературная. Формат 60×90¹/₁₆.

Печать офсетная. Усл. п. л. 25,00. Тираж 1500 экз.

Заказ № .

Отпечатано в полном соответствии с качеством
предоставленных материалов в ОАО «Дом печати — ВЯТКА»
610033, г. Киров, ул. Московская, 122