

К. В. ШАЛИМОВА

ФИЗИКА ПОЛУПРОВОДНИКОВ

*Допущено Министерством высшего
и среднего специального образования СССР
в качестве учебного пособия для студентов вузов*



«Энергия» · Москва · 1971

6Ф0.32

Ш18

УДК 621.315.592/075/53

Ш18

К. В. Шалимова

Физика полупроводников. Учебное пособие для вузов, М., «Энергия», 1971.

312 с. с илл.

В книге изложены основные вопросы по физике полупроводников. В ней рассмотрена элементарная теория проводимости и модельные представления о механизме проводимости полупроводников, даны основы зонной теории полупроводников, изложена статистика электронов и дырок, рассмотрены механизмы рассеяния носителей заряда, генерация и рекомбинация носителей, диффузия и дрейф неравновесных носителей заряда, а также изложены контактные и поверхностные явления в полупроводниках, оптические и фотоэлектрические свойства полупроводников.

Книга предназначена для студентов вузов и может быть полезна инженерно-техническим работникам.

3-3-12

6Ф0.32

260-71

О Г Л А В Л Е Н И Е

Предисловие	6
Наиболее часто встречающиеся обозначения	7

Глава первая. Элементарная теория проводимости и эффект Холла

1-1. Классификация веществ по величине проводимости	9
1-2. Модельные представления о механизме проводимости полупроводников	14
1-3. Проводимость полупроводника	21
1-4. Эффект Холла	23
1-5. Эффект Холла для полупроводника со смешанной проводимостью	26

Глава вторая. Основы зонной теории полупроводников

2-1. Уравнение Шредингера для кристалла	28
2-2. Адиабатическое приближение и валентная аппроксимация	30
2-3. Одноэлектронное приближение	31
2-4. Приближение сильно связанных электронов	33
2-5. Число состояний в разрешенной зоне	38
2-6. Зоны Бриллюэна	42
2-7. Зависимость энергии электронов от волнового вектора у дна и потолка энергетической зоны	44
2-8. Движение электронов в кристалле под действием внешнего электрического поля	47
2-9. Эффективная масса носителя заряда	51
2-10. Структура энергетических зон некоторых полупроводников	57
2-11. Элементарная теория локальных уровней.	61

Глава третья. Статистика электронов и дырок в полупроводниках

3-1. Плотность квантовых состояний	66
3-2. Функция распределения Ферми — Дирака.	70
3-3. Степень заполнения примесных уровней	72
3-4. Концентрация электронов и дырок	74
3-5. Примесный полупроводник	77
3-6. Собственный полупроводник	82
3-7. Зависимость положения уровня Ферми от концентрации примеси и температуры для невырожденного полупроводника	86
3-8. Зависимость положения уровня Ферми от температуры для невырожденного полупроводника с частично компенсированной примесью	93
3-9. Примесные зоны	96

Глава четвертая. Механизмы рассеяния носителей заряда в полупроводниках

4-1. Кинетическое уравнение Больцмана	100
4-2. Равновесное состояние	105
4-3. Время релаксации	105

4-4. Эффективное сечение рассеяния. Типы центров рассеяния	109
4-5. Рассеяние на ионах примеси	111
4-6. Рассеяние на атомах примеси и дислокациях	115
4-7. Тепловые колебания атомов одномерной решетки	116
4-8. Тепловые колебания атомов трехмерной решетки. Статистика фононов	126
4-9. Рассеяние на тепловых колебаниях решетки	132

Глава пятая. Кинетические явления в полупроводниках

5-1. Электропроводность полупроводников	139
5-2. Зависимость подвижности носителей заряда от температуры	141
5-3. Эффект Холла	146
5-4. Термоэлектрические явления	151
5-5. Явления переноса в сильных электрических полях	158

Глава шестая. Генерация и рекомбинация неравновесных носителей заряда

6-1. Равновесные и неравновесные носители заряда	169
6-2. Биполярная световая генерация носителей заряда	172
6-3. Монополярная световая генерация. Максвелловское время релаксации	174
6-4. Виды рекомбинации	175
6-5. Межзонная излучательная рекомбинация	176
6-6. Межзонная ударная рекомбинация	181
6-7. Рекомбинация носителей заряда через центры захвата	183
6-8. Температурная зависимость времени жизни носителей заряда при рекомбинации через локальные уровни	190
6-9. Центры прилипания и центры рекомбинации	192

Глава седьмая. Диффузия и дрейф неравновесных носителей заряда

7-1. Уравнение непрерывности	195
7-2. Диффузионный и дрейфовый токи	197
7-3. Соотношение Эйнштейна	199
7-4. Диффузия и дрейф неравновесных носителей заряда в случае монополярной проводимости	200
7-5. Движение неосновных носителей заряда	202
7-6. Биполярный коэффициент диффузии и биполярная дрейфовая подвижность	207

Глава восьмая. Контактные явления в полупроводниках

8-1. Полупроводник во внешнем электрическом поле	210
8-2. Работа выхода	214
8-3. Контакт металл — металл. Контактная разность потенциалов	216
8-4. Контакт металл — полупроводник	219
8-5. Выпрямление на контакте металл — полупроводник	222
8-6. Диодная теория выпрямления	225
8-7. Диффузионная теория выпрямления	228
8-8. Контакт электронного и дырочного полупроводников	230
8-9. Выпрямление на <i>p-n</i> переходе	234
8-10. Теория тонкого <i>p-n</i> перехода	236

Глава девятая. Поверхностные явления в полупроводниках

9-1. Поверхностный потенциал	242
9-2. Эффект поля	246

9-3. Поверхностная рекомбинация	251
9-4. Влияние поверхностной рекомбинации на распределение неравно- весных носителей заряда	254

Глава десятая. Оптические свойства полупроводников

10-1. Спектр поглощения света	256
10-2. Собственное поглощение света при прямых переходах	257
10-3. Собственное поглощение света при непрямах переходах	262
10-4. Экситонное поглощение света	267
10-5. Поглощение света свободными носителями заряда	269
10-6. Примесное поглощение	273
10-7. Поглощение света решеткой	275
10-8. Зависимость положения края собственного поглощения от темпера- туры и давления	275
10-9. Спонтанное и вынужденное излучение атомов	276
10-10. Люминесценция полупроводников	281
10-11. Мономолекулярное свечение твердых тел	282
10-12. Рекомбинационное свечение полупроводников	283
10-13. Релаксация фотолюминесценции полупроводников	285
10-14. Температурное тушение люминесценции полупроводников	287
10-15. Стимулированное излучение в полупроводниках	288

Глава одиннадцатая. Фотоэлектрические явления в полупроводниках

11-1. Внутренний фотоэффект	291
11-2. Фотопроводимость	292
11-3. Релаксация фотопроводимости	294
11-4. Фотопроводимость при наличии поверхностной рекомбинации и диф- фузии носителей заряда	298
11-5. Примесная фотопроводимость	301
11-6. Эффект Дембера	302
11-7. Фотомагнитноэлектрический эффект	304
11-8. Вентильная фотоэлектродвижущая сила	308

ПРЕДИСЛОВИЕ

При написании данного учебного пособия автор исходил из следующих соображений.

Учебное пособие должно отвечать задаче подготовки студентов, специализирующихся в области полупроводниковой техники. Поэтому наряду с теоретическим характером материала, обусловливающим высокую математическую насыщенность, должно быть уделено большое внимание физической интерпретации явлений и экспериментальным результатам их исследований на конкретных полупроводниковых веществах. При этом, кроме разбора общих фундаментальных проблем физики полупроводников, необходимо и рассмотрение ряда явлений прикладного характера, играющих большую роль для дальнейшего понимания процессов, происходящих в полупроводниковых приборах.

Этими положениями и определялся отбор материала для настоящего учебного пособия, написанного на основе лекций по физике полупроводников, читаемых автором в течение более пятнадцати лет для студентов специальности «Полупроводниковые приборы» в Московском ордена Ленина энергетическом институте.

Многолетний педагогический опыт автора показывает, что перегруженность математическим аппаратом материала, предназначенного не для физиков-теоретиков, отрицательно сказывается на прочном усвоении студентами физической стороны обсуждаемых явлений. Отнюдь не ратуя за излишнее упрощенчество, в местах, где строгость изложения не страдала бы от отсутствия сложных математических выкладок, автор старался их избежать, давая ссылки на оригинальные работы.

Автор выражает большую благодарность профессорам В. В. Пасынкову и В. Л. Конькову за ценные советы и замечания, сделанные ими при рецензировании рукописи. Автор выражает признательность сотрудникам и аспирантам кафедры «Полупроводниковые приборы» МЭИ за помощь в работе над созданием данного пособия.

Все замечания и предложения читателей просьба направлять по адресу: Москва, Е-250, Красноказарменная 14, МЭИ, кафедра «Полупроводниковые приборы».

НАИБОЛЕЕ ЧАСТО ВСТРЕЧАЮЩИЕСЯ ОБОЗНАЧЕНИЯ

- ρ — удельное сопротивление, линейная плотность, объемный заряд;
 σ — удельная проводимость, эффективное сечение;
 α — коэффициент поглощения, термический коэффициент сопротивления, термо-э. д. с., коэффициент ударной ионизации;
 T — абсолютная температура, время, кинетическая энергия, период;
 l — средняя длина пробега, длина;
 τ — среднее время пробега, время жизни, время релаксации;
 v_0 — средняя скорость теплового движения;
 E — энергия носителей заряда;
 e — заряд электрона, величина деформации;
 c — скорость света;
 C — емкость;
 ΔE_g — ширина запрещенной зоны;
 ΔE_{g0} — ширина запрещенной зоны при 0°K ;
 a — постоянная решетки, постоянная величина, ускорение;
 R — коэффициент Холла, коэффициент отражения;
 μ , μ_d — дрейфовая подвижность носителей заряда;
 μ_H — холловская подвижность носителей заряда;
 φ — угол Холла, угол, функция;
 h — постоянная Планка, $\hbar = h/2\pi$;
 n_i — собственная концентрация электронов, дырок;
 g — радиус-вектор электрона;
 R — радиус-вектор ядра;
 k — волновой вектор носителя заряда;
 p — квазиимпульс носителя заряда;
 k_0 — постоянная Больцмана;
 V_B — объем зоны Бриллюэна;
 ω — циклическая частота;
 a, b, c — полуоси эллипсоида, постоянные величины;
 m, M — масса свободного электрона и атома (иона) соответственно;
 m^* — эффективная масса носителя заряда;
 m_n^*, m_p^* — эффективная масса электрона и дырки соответственно;
 ω_c — циклотронная частота;
 i, j, k — единичные векторы;
 Å — ангстрем (10^{-10} м);
 λ — длина волны;
 dZ — число квантовых состояний;
 N — плотность квантовых состояний, концентрация;
 F — уровень Ферми (энергия Ферми), сила, функция;
 f_0, f_{0p} — функция распределения при тепловом равновесии электронов и дырок соответственно;
 f — неравновесная функция распределения;
 N_d, N_a — концентрация донорной и акцепторной примеси соответственно;
 n_d, n_a — концентрация электронов на донорной и акцепторной примеси соответственно;
 E_d, E_a — энергия донорного и акцепторного уровней соответственно;
 \mathcal{E}, \mathcal{E} — напряженность электрического поля;
 \mathbf{H} — напряженность магнитного поля, оператор Гамильтона;
 \mathbf{B} — магнитная индукция;
 \mathbf{v}, v — скорость движения носителя заряда;
 $v_{гр}, v_\phi, v_{зв}$ — групповая скорость, фазовая скорость и скорость звука соответственно;
 κ — относительная магнитная проницаемость вещества;
 κ_0 — магнитная постоянная;

- ϵ — диэлектрическая проницаемость материала;
 ϵ_0 — диэлектрическая постоянная;
 V — потенциал кристаллической решетки, разность потенциалов, объем кристалла;
 j, j — плотность тока;
 I — ток, интенсивность светового потока;
 n — концентрация электронов, главное квантовое число;
 p, P — концентрация дырок, показатель степени, давление, вероятность перехода, поляризуемость среды;
 p_d, p_a — концентрация ионизированной донорной и акцепторной примеси;
 β — фактор спинового вырождения, квантовый выход, температурный коэффициент, коэффициент квазиупругой силы;
 η — приведенный уровень Ферми, импульс фотона;
 W — энергия активации, вероятность;
 δ — относительное изменение энергии, дельта-функция, толщина туннельного барьера;
 q, q — волновой вектор фонона;
 Π — коэффициент Пельтье;
 τ_T — коэффициент Томсона;
 S, s — энтропия, путь, площадь, скорость поверхностной рекомбинации, функция;
 U — потенциальная энергия, контактная разность потенциалов;
 j_s — плотность тока насыщения;
 v_n — скорость движения домена;
 \mathcal{L} — длина образца, толщина слоя объемного заряда;
 E_n — энергия ионизации;
 A, B, V, C — постоянные величины;
 g — скорость генерации носителей заряда, кратность уровня, вероятность поглощения;
 ξ — химический потенциал;
 r — скорость рекомбинации носителей заряда, коэффициент пропорциональности;
 γ — коэффициент рекомбинации, коэффициент пропорциональности;
 F_n, F_p — квазиуровень Ферми электронов и дырок соответственно;
 $\Delta n, \Delta p$ — концентрация неравновесных электронов и дырок соответственно;
 Q — квазиимпульс решетки, объемный заряд;
 R_n^0, R_p^0 — плотность равновесного излучения и поглощения соответственно;
 Ψ — функция, поверхностный потенциал;
 ν — частота, постоянная спада;
 \bar{n} — показатель преломления среды;
 L — диффузионная длина носителя заряда, длина затягивания, длина;
 Φ — функция, работа выхода;
 χ — работа выхода, показатель поглощения;
 D — коэффициент диффузии носителей заряда, прозрачность потенциального барьера (постоянная величина);
 t — время;
 u — величина смещения, плотность излучения;
 θ — угол, температура Дебая;
 ϕ — угол;
 k_p, k_n — вероятность захвата носителя противоположного знака.
 \hat{H} — оператор Гамельтона.

Глава первая

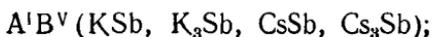
ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ПРОВОДИМОСТИ И ЭФФЕКТ ХОЛЛА

1-1. КЛАССИФИКАЦИЯ ВЕЩЕСТВ ПО ВЕЛИЧИНЕ ПРОВОДИМОСТИ

Все вещества в природе по своим электрофизическим свойствам могут быть разделены на три больших класса: металлы, полупроводники и диэлектрики. Наиболее проста классификация веществ по их удельному электрическому сопротивлению. У металлов эта величина лежит в пределах от 10^{-6} до 10^{-4} *ом·см* (например, удельное сопротивление серебра при комнатной температуре равно $1,58 \cdot 10^{-6}$ *ом·см*; сплав нихрома имеет удельное сопротивление $1,05 \cdot 10^{-4}$ *ом·см*). Группа веществ с удельным сопротивлением от 10^{-4} до 10^{10} *ом·см* может быть отнесена 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*, а германия от $5 \cdot 10^{-4}$ до 47 *ом·см*). Наконец, вещества с удельным сопротивлением более 10^{10} *ом·см* относятся к классу диэлектриков (например, при 200° С удельное сопротивление слюды в зависимости от ее состава составляет величину 10^{13} — 10^{16} *ом·см*, стекла — 10^8 — 10^{15} *ом·см*).

Из приведенных данных видно, что резкого изменения в величине удельного сопротивления при переходе от одного класса веществ к другому не существует. Для полупроводников и диэлектриков это отражает их принципиальное качественное сходство. Более различны по своей природе металлы и полупроводники.

Полупроводниками являются как простые химические элементы, такие как В, С, Si, Ge, α — Sn, P, As, β — Sb, α — S, Se (красный), Te, I, так и многие сложные химические соединения. В качестве примеров ниже приведены некоторые соединения элементов периодической системы Менделеева (рис. 1-1), обладающие полупроводниковыми свойствами. К ним относятся бинарные соединения типа:



ПЕРИОДИЧЕСКАЯ СИСТЕМА

ПЕРИОДЫ	РЯДЫ	ГРУППЫ					
		I	II	III	IV	V	
1	I	H ¹ ВОДОРОД 1,00797					
2	II	Li ³ ЛИТИЙ 6,939	Be ⁴ БЕРИЛЛИЙ 9,0122	B ⁵ БОР 10,811	C ⁶ УГЛЕРОД 12,01115	N ⁷ АЗОТ 14,0067	
3	III	Na ¹¹ НАТРИЙ 22,9898	Mg ¹² МАГНИЙ 24,315	Al ¹³ АЛЮМИНИЙ 26,9815	Si ¹⁴ КРЕМНИЙ 28,086	P ¹⁵ ФОСФОР 30,9738	
4	IV	K ¹⁹ КАЛИЙ 39,102	Ca ²⁰ КАЛЬЦИЙ 40,08	Sc ²¹ СКАНДИЙ 44,956	Ti ²² ТИТАН 47,90	V ²³ ВАНАДИЙ 50,942	
	V	Cu ²⁹ МЕДЬ 63,546	Zn ³⁰ ЦИНК 65,37	Ga ³¹ ГАЛЛИЙ 69,72	Ge ³² ГЕРМАНИЙ 72,59	As ³³ МЫШЬЯК 74,9216	
5	VI	Rb ³⁷ РУБИДИЙ 85,47	Sr ³⁸ СТРОНЦИЙ 87,62	Y ³⁹ ИТРИЙ 88,905	Zr ⁴⁰ ЦИРКОНИЙ 91,22	Nb ⁴¹ НИОБИЙ 92,906	
	VII	Ag ⁴⁷ СЕРЕБРО 107,868	Cd ⁴⁸ КАДМИЙ 112,40	In ⁴⁹ ИНДИЙ 114,82	Sn ⁵⁰ ОЛОВО 118,69	Sb ⁵¹ СУРЬМА 121,75	
6	VIII	Cs ⁵⁵ ЦЕЗИЙ 132,905	Ba ⁵⁶ БАРИЙ 137,34	La * ЛАНТАН 138,91	Hf ⁷² ГАФНИЙ 178,49	Ta ⁷³ ТАНТАЛ 180,948	
	IX	Au ⁷⁹ ЗОЛОТО 196,967	Hg ⁸⁰ РТУТЬ 200,59	Tl ⁸¹ ТАЛЛИЙ 204,37	Pb ⁸² СВИНЕЦ 207,19	Bi ⁸³ ВИСМУТ 208,980	
7	X	Fr ⁸⁷ ФРАНЦИЙ [223]	Ra ⁸⁸ РАДИЙ [226]	Ac ** АКТИНИЙ [227]	Ku ¹⁰⁴ КУРЧАТОВИЙ [260]		
* Л А Н Т А Н							
	58	59	60	61	62	63	64
	Ce ² ЦЕРИЙ 140,12	Pr ² ПРАЗЕОДИМ 140,907	Nd ² НЕОДИМ 144,24	Pm ² ПРОМЕТИЙ [147]	Sm ² САМАРИЙ 150,35	Eu ² ЕВРОПИЙ 151,96	Gd ² ГАДОЛИНИЙ 157,25
** А К Т И							
	90	91	92	93	94	95	96
	Th ² ТОРИЙ 232,038	Pa ² ПРОТАКТИНИЙ [231]	U ² УРАН 238,03	Np ² НЕПТУНИЙ [237]	Pu ² ПЛУТОНИЙ [244]	Am ² АМЕРИЦИЙ [243]	Cm ² КЮРИЙ [247]

Рис. 1-1. Периодическая система

ЭЛЕМЕНТОВ Д.И. МЕНДЕЛЕЕВА

Э Л Е М Е Н Т О В										ЭЛЕКТРОННЫЕ СЛОИ	
VI	VII		VIII				0				
	(H)						He ² ГЕЛИЙ 4,0026		1	K	
8 O КИСЛОРОД 15,9994	9 F ФТОР 18,9984					Ne ¹⁰ НЕОН 20,179		8 2	L K		
16 S СЕРА 32,064	17 Cl ХЛОР 35,453					Ar ¹⁸ АРГОН 39,948		8 8 2	M L K		
24 Cr ХРОМ 51,996	25 Mn МАРГАНЕЦ 54,9380	26 Fe ЖЕЛЕЗО 55,847	27 Co КОБАЛЬТ 58,9332	28 Ni НИКЕЛЬ 58,71						N M L K	
34 Se СЕЛЕН 78,96	35 Br БРОМ 79,904					Kr ³⁶ КРИПТОН 83,80		8 18 8 2	N M L K		
42 Mo МОЛИБДЕН 95,94	43 Tc ТЕХНЕЦИЙ [99]	44 Ru РУТЕНИЙ 101,07	45 Rh РОДИЙ 102,905	46 Pd ПАЛЛАДИЙ 106,4						O N M L K	
52 Te ТЕЛУР 127,60	53 I ИОД 126,9044					Xe ⁵⁴ КСЕНОН 131,30		8 18 18 8 2	O N M L K		
74 W ВОЛЬФРАМ 183,85	75 Re РЕНИЙ 186,2	76 Os ОСМИЙ 190,2	77 Ir ИРИДИЙ 192,2	78 Pt ПЛАТИНА 195,09						P O N M L K	
84 Po ПОЛОНИЙ [210]	85 At АСТАТ [210]					Rn ⁸⁶ РАДОН [222]		8 18 32 18 8 2	P O N M L K		

О И Д Ы 58-71

65 Tь ТЕРБИЙ 158,924	66 Dy ДИСПРОЗИЙ 162,50	67 Ho ГОЛЬМИЙ 164,930	68 Er ЭРБИЙ 167,26	69 Tu ТУЛИЙ 168,934	70 Yь ИТТЕРБИЙ 173,04	71 Lu ЛЮТЕЦИЙ 174,97	P O N M L K
----------------------------	------------------------------	-----------------------------	--------------------------	---------------------------	-----------------------------	----------------------------	----------------------------

Н О И Д Ы 90-103

97 Bк БЕРКЛИЙ [247]	98 Cf КАЛИФОРНИЙ [252]	99 Es ЭЙНШТЕЙНИЙ [254]	100 Fm ФЕРМИЙ [257]	101 Md МЕНДЕЛЕВИЙ [257]	102 (No) (НОБЕЛИЙ) [255]	103 (Lr) (ЛОУРЕНСИЙ) [256]	O P O N M L K
---------------------------	------------------------------	------------------------------	---------------------------	-------------------------------	--------------------------------	----------------------------------	---------------------------------

элементов Менделеева.

$A^{II}B^{VII}$ (ZnCl₂, ZnI₂, CdCl₂, CdI₂);
 $A^{II}B^{VI}$ (ZnO, ZnS, ZnSe, ZnTe, CdO, CdS, CdSe, CdTe, HgS, HgSe, HgTe);
 $A^{II}B^V$ (ZnSb, CdSb, Mg₃Sb₂, Zn₃As₂, Cd₃P₂, Cd₃As₂);
 $A^{II}B^{IV}$ (Mg₂Si, Mg₂Ge, Mg₂Sn, Ca₂Si, Ca₂Sn, Ca₂Pb);
 $A^{III}B^{VI}$ (GaS, GaSe, InS, InSe, In₂O₃, In₂S₃, In₂Se₃, In₂Te₃, Tl₂S);
 $A^{III}B^V$ (AlP, AlSb, AlAs, GaP, GaSb, GaAs, InP, InSb, InAs);
 $A^{IV}B^{VI}$ (GeO₂, PbS, PbSe, PbTe, TiO₂, GeTe, SnTe, GeS);
 $A^{IV}B^{IV}$ (SiC, SiGe);
 A^VB^{VI} (As₂S₃, As₂Se₃, Sb₂Se₃, Sb₂Te₃, Bi₂S₃, Bi₂Se₃, Bi₂Te₃);
 $A^{VI}B^{VI}$ (MoO₃, WO₃);
 $A^{VIII}B^{VI}$ (Fe₂O₃, NiO);

тройные соединения типа:

$A^I B^{III} B_2^{VI}$ (CuAlS₂, CuInS₂, CuInSe₂, CuInTe₂, AgInSe₂, AgInTe₂, CuGaSe₂, CuGaTe₂);
 $A^I B^V B_2^{VI}$ (CuSbS₂, CuAsS₂, AgSbSe₂, AgSbTe₂, AgBiS₂, AgBiSe₂, AgBiTe₂);
 $A^I B^{VIII} B_2^{VI}$ (CuFeSe₂, AgFeSe₂, AgFeTe₂);
 $A^{II} B^{IV} B_2^V$ (ZnSiAs₂, ZnGeAs₂);
 $A^{IV} B^V B^{VI}$ (PbBiSe₂).

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

При классификации веществ по величине удельного сопротивления трудно разделить полупроводники и металлы. Однако при снятии температурных зависимостей удельного электрического сопротивления это различие часто проявляется достаточно четко. С увеличением температуры у некоторых химически чистых металлов удельное сопротивление ρ примерно пропорционально абсолютной температуре T , т. е.

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

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

Графики температурной зависимости удельного сопротивления некоторых металлов представлены на рис. 1-2, а.

Характерным для температурной зависимости удельной проводимости металлов, как следует из рис. 1-2, б, является наличие отрицательного температурного коэффициента электропроводности, т. е. удельная проводимость металла σ_1 при температуре T_1 больше его проводимости σ_2 при более высокой температуре T_2 .

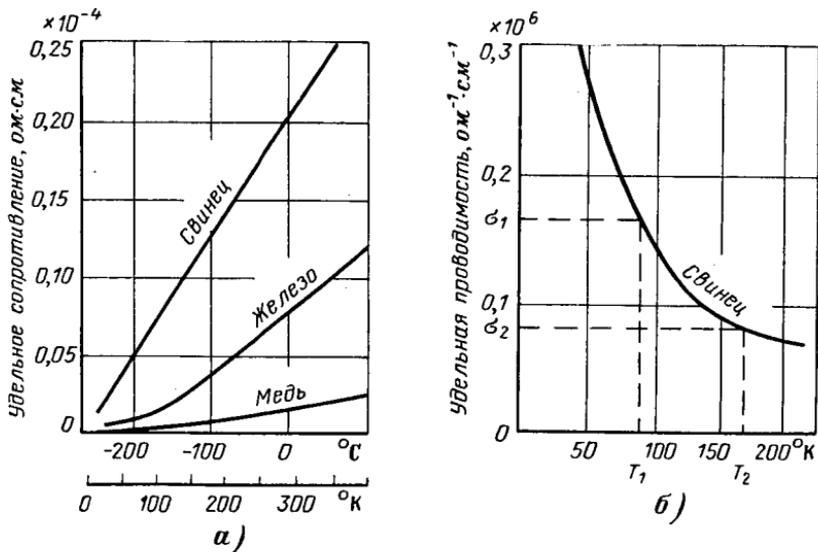


Рис. 1-2. Изменение удельного сопротивления чистых металлов (а) и удельной проводимости свинца (б) в зависимости от температуры.

Для полупроводников характер температурной зависимости удельного сопротивления и проводимости иной:

$$\rho = \rho_0 e^{\beta/T}; \tag{1-2}$$

$$\sigma = \sigma_0 e^{-\frac{\beta}{T}}, \tag{1-2a}$$

где ρ_0 , σ_0 , β — некоторые положительные по знаку константы;
 T — температура, $^{\circ}\text{K}$.

Из графиков температурной зависимости удельной проводимости полупроводников, которые изображены на рис. 1-3, видно, что для них температурный коэффициент положителен, т. е. при $T_2 > T_1$ $\Delta\sigma = \sigma_2 - \sigma_1 > 0$.

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 при определенных условиях полупроводники могут вести себя подобно металлам. Поэтому знак температурного коэффициента проводимости не всегда может однозначно охарактеризовать природу вещества.

Ответить на этот вопрос можно, если проследить, как меняется сопротивление вещества электрическому току при понижении температуры. Как следует из рис. 1-2, б, с понижением температуры проводимость металлов растет. Это возможно потому, что независимо от температуры в металле всегда имеются свободные

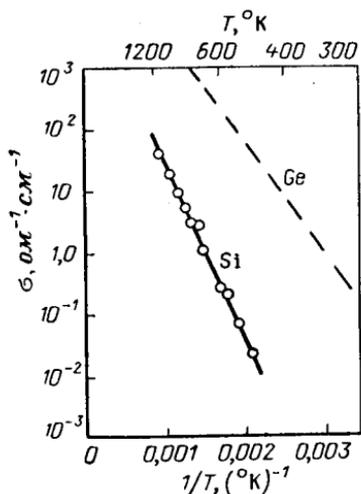


Рис. 1-3. Температурная зависимость удельной проводимости полупроводника.

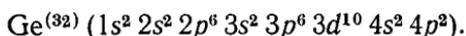
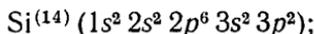
носители заряда — электроны, для образования которых не требуется дополнительной энергии. У полупроводников проводимость уменьшается при понижении температуры (рис. 1-3), а при приближении температуры к абсолютному нулю полупроводник при определенных условиях ведет себя как диэлектрик. Следовательно, для образования в полупроводнике свободных носителей заряда необходимо сообщить ему энергию. При этом, как видно из соотношения (1-2а), с увеличением температуры нагрева полупроводника имеет место быстрый рост числа носителей заряда.

Естественно, что сообщать энергию полупроводнику для создания в нем свободных носителей заряда можно не только нагреванием, но и другими методами. Свободные носители

заряда появляются при воздействии на полупроводник различных видов радиации (света, ядерного излучения и т. п.), сильных электрических и магнитных полей и др.

1-2. МОДЕЛЬНЫЕ ПРЕДСТАВЛЕНИЯ О МЕХАНИЗМЕ ПРОВОДИМОСТИ ПОЛУПРОВОДНИКОВ

Рассмотрим механизм проводимости типичных полупроводников: кремния Si и германия Ge. Эти элементы находятся в одной подгруппе периодической таблицы Менделеева с углеродом (рис. 1-1) и имеют кристаллическую решетку типа алмаза (рис. 1-4). Следовательно, заполнение внешних электронных оболочек у них одинаково. Распределение электронов по состояниям в атоме кремния и германия следующее:



Наивысшая валентность их равна четырем. При образовании твердого состояния 3s и 3p электроны атомов кремния (4s и 4p атомов германия) имеют неспаренные спины, они и участвуют в образовании ковалентной связи. Схематически кристаллическую решетку

с парноэлектронной связью можно представить, как изображено на рис. 1-5, а. Здесь в узле решетки находится атомный остаток с зарядом $+4$, которому принадлежат четыре валентных электрона. Валентные электроны, образующие парную связь, на рис. 1-5 изображены черными точками.

В целом система, представленная на рис. 1-5, а, электронейтральна. Если поместить ее в электрическое поле, то электрический ток не возникнет, так как все связи в решетке заполнены и свободных носителей заряда нет.

Допустим, что под воздействием каких-то возмущений произошел разрыв валентной связи и электрон стал свободным. В результате образовалась ненасыщенная связь, и в этом месте как бы сосредоточился положительный заряд. Такая незавершенная валентная связь получила название дырки. Незавершенная валентная

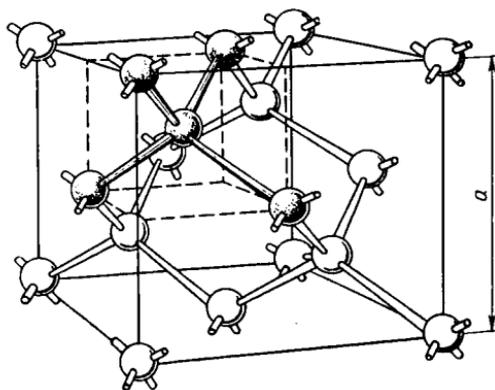


Рис. 1-4. Кристаллическая решетка типа алмаза. a — постоянная решетки.

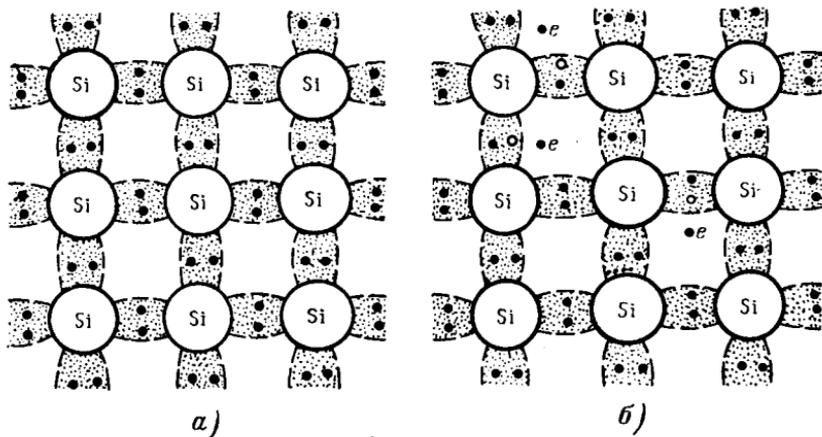


Рис. 1-5. Двумерное представление расположения связей в решетке кремния (собственный полупроводник).

связь может быть заполнена электроном, перешедшим на нее с соседней насыщенной связью, а следовательно, она будет перемещаться по кристаллу вследствие обмена электронами между атомами. Однако в целом система остается электронейтральной, 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 точками.

Полупроводник, в котором в результате разрыва валентных связей образуется равное количество свободных электронов и незавершенных связей — дырок, называется собственным. При комнатной температуре в германии концентрация свободных электронов и дырок составляет 10^{13} см^{-3} , а у кремния около 10^{10} см^{-3} . Увеличение температуры, естественно, приводит к возрастанию числа разорванных валентных связей, а следовательно, к росту концентрации свободных носителей заряда в полупроводнике.

Свободные носители за счет тепловой энергии перемещаются по кристаллу. Но в реальном веществе всегда есть нарушения кристаллической структуры, обусловленные тепловыми колебаниями атомов кристалла, наличием разного вида дефектов, примесей, дислокаций. Поэтому свободный носитель заряда, перемещаясь по кристаллу, будет сталкиваться с неоднородностями кристаллической

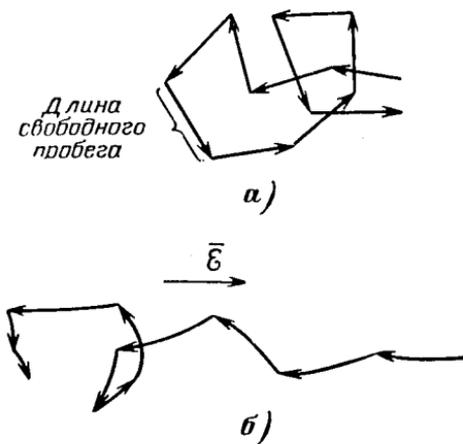


Рис. 1-6. Схема движения свободного электрона за счет тепловой энергии (а) и во внешнем электрическом поле (б).

решетки, в результате чего меняется направление его движения. В силу этого тепловое движение свободного носителя заряда является беспорядочным, как это изображено на рис. 1-6, а. Расстояние, проходимое, например, электроном между двумя столкновениями, называется длиной свободного пробега, а усредненное значение всех отрезков пути есть средняя длина свободного пробега. Соответственно время между двумя соударениями и его усредненное значение называется временем свободного пробега и его средним значением.

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

$$l = v_0 \tau, \quad (1-3)$$

где v_0 — средняя скорость теплового движения электронов (среднее значение модуля скорости). При комнатной температуре она составляет величину порядка 10^7 см/сек .

Поскольку движение носителей заряда имеет хаотический характер, то вектор их средней скорости равен нулю, т. е. каждому свободному носителю заряда будет соответствовать носи-

тель со скоростью, вектор которой направлен в обратную сторону. Схематически это представлено на рис. 1-7, а. В силу этого средняя скорость направленного движения всех носителей заряда равна нулю, поэтому нет переноса заряда.

Поместим собственный полупроводник в электрическое поле. Его свободные электроны под действием электрического поля начнут двигаться в направлении, противоположном направлению поля. Те электроны, направление движения которых за счет тепловой энергии было против направления электрического поля, будут ускоряться полем и за время свободного пробега будут накапливать энергию. Электроны, движущиеся по полю, будут замедляться им, и их энергия будет уменьшаться. Это иллюстрируется рис. 1-6, б и 1-7, б. Такие электроны при столкновении с неоднородностями кристаллической решетки будут либо отдавать ей излишек энергии, приобретенный под воздействием внешнего поля, либо получать от нее энергию, в результате чего они придут в тепловое равновесие с решеткой. После столкновения электроны опять будут иметь направленное движение во внешнем электрическом поле. *Направленное движение свободных носителей заряда в электрическом поле носит название дрейфа, а скорость их направленного движения называется дрейфовой скоростью.* На рис. 1-7, б жирными линиями показаны векторы дрейфовой составляющей скорости, а пунктиром — векторы суммарной скорости теплового движения и дрейфа.

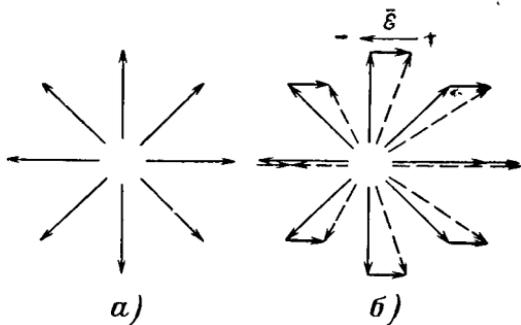


Рис. 1-7. Схематическое изображение скоростей электронов проводимости при отсутствии (а) и наличии (б) электрического поля.

В этом случае средняя скорость движения свободных электронов уже не равна нулю (рис. 1-7, б), и через полупроводник в направлении приложенного поля потечет ток, обусловленный свободными электронами.

Электроны в насыщенных связях при переходе на незавершенную связь под действием электрического поля будут перемещаться в направлении против направления поля. Тем самым вакантное место в валентной связи будет также перемещаться, но уже по направлению поля, что равносильно перемещению по полю положительного заряда. Механизм проводимости, обусловленный движением связанных электронов по вакантным связям, получил название дырочной проводимости.

Следовательно, электрический ток в собственном полупроводнике определяется двумя составляющими — электронным и дырочным токами, текущими в одном направлении.

Проводимость собственного полупроводника можно объяснить, если исходить из энергетических соображений. В твердом теле энергетическое состояние свободных электронов отлично от состояния связанных электронов. Это можно охарактеризовать с помощью энергетической диаграммы, как представлено на рис. 1-8.

Здесь E_c — минимальная энергия, которую может иметь свободный электрон. Возможные значения энергии свободных электронов образуют зону, которая называется свободной или зоной проводимости. Величина E_v является максимальной энергией электронов в полностью завершенной связи. Ниже E_v лежит спектр энергий всех связанных валентных электронов, и эта зона называется заполненной или валентной зоной. Если все

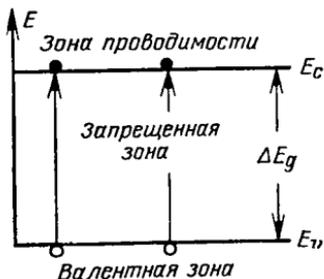


Рис. 1-8. Схематическое изображение энергетических зон собственного полупроводника.

связи завершены, то все состояния в валентной зоне заняты, и в зоне проводимости нет электронов.

Для того чтобы разорвать парно-электронную связь, необходимо затратить определенную энергию. Величина этой энергии ΔE_g на рис. 1-8 характеризует ширину запрещенной зоны полупроводника. Разрыв парноэлектронной связи в энергетической диаграмме равнозначен переводу электрона из нижней зоны в верхнюю, в результате чего появляется электрон в зоне проводимости и свободная дырка в валентной зоне.

В этом случае свободный электрон обладает энергией, большей той, которую он имел в связанном состоянии, на величину не менее энергии ширины запрещенной зоны.

Ширина запрещенной зоны определяется природой химической связи, а также зависит от сорта атомов, образующих кристаллическую решетку. В кристаллах, состоящих из элементов IV группы периодической системы Менделеева, она убывает по мере роста атомного номера элемента. Это можно проиллюстрировать данными табл. 1-1.

Из табл. 1-1 следует, что решетка алмаза более плотная и, следовательно, более прочная. Электронные облака больше перекрываются в плотной решетке, и на разрыв парноэлектронной связи требуется больше энергии, чем для более рыхлой связи.

Рассмотрим теперь модель механизма проводимости полупроводника с решеткой типа алмаза, в котором один из атомов замещен атомом элемента V группы, например, мышьяк в решетке кремния. *Полупроводник, имеющий примеси, называется примесным, а проводимость, созданная введенной примесью, носит название примесной проводимости.*

Пять валентных электронов атома мышьяка 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 че-

тырьмя электронами ближайших атомов кремния объединены в парноэлектронные связи, как это схематически представлено на

Т а б л и ц а 1

Экспериментальные значения параметров решетки и ширины запрещенной зоны для кристаллов типа алмаза

Элемент	Постоянная решетки, Å	Число атомов в 1 см ³	Ширина запрещенной зоны (при 0° К), эВ
Алмаз	3,56	$17,7 \cdot 10^{22}$	7,0
Кремний	5,42	$5 \cdot 10^{22}$	1,14
Германий	5,62	$4,52 \cdot 10^{22}$	0,75
Серое олово	6,46	$2,82 \cdot 10^{22}$	0,08

рис. 1-9, а. Пятый электрон слабо связан с атомом мышьяка, но при низких температурах локализован около него. При повышении температуры пятый электрон атома мышьяка будет оторван от примеси и может свободно перемещаться по кристаллу. В этом

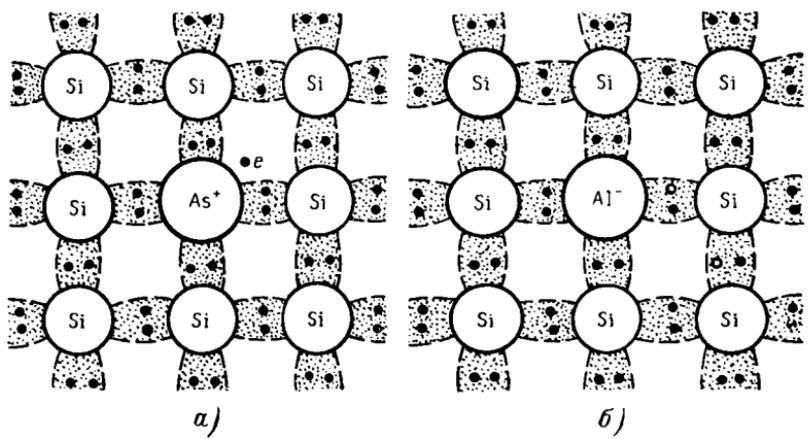


Рис. 1-9. Схематическое изображение кристаллической решетки донорного (а) и акцепторного (б) полупроводников.

случае также соблюдается электронейтральность кристалла, так как атом мышьяка, отдавший пятый электрон, будет 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 проводимость,

количество электронов, оторванных от примеси, будет значительно больше количества электронов и дырок, образовавшихся в результате разрыва валентных связей. В силу этого *доминирующую роль в проводимости кристалла будут играть электроны, и поэтому они называются основными носителями заряда, а дырки — неосновными носителями заряда. Такой полупроводник называется электронным или n-типа, а примесь, отдающая электроны, носит название донорной.*

На энергетической диаграмме наличие примеси в решетке полупроводника будет характеризоваться появлением локального уровня, лежащего в запрещенной зоне. Так как при ионизации атома мышьяка образуется свободный электрон и для его отрыва требуется значительно меньшая энергия, чем для разрыва валентных связей кремния, то энергетический уровень донорной примеси

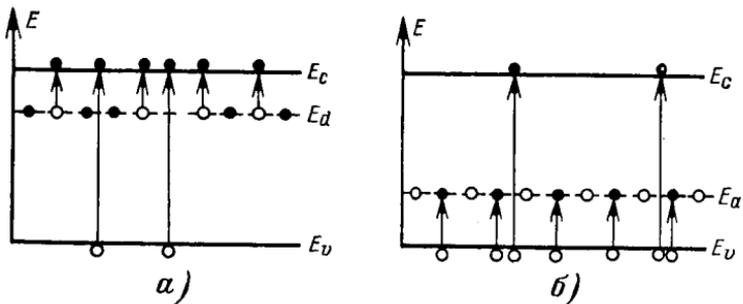


Рис. 1-10. Энергетическая диаграмма донорного (а) и акцепторного (б) полупроводников.

E_d должен располагаться в запрещенной зоне на небольшой глубине под дном зоны проводимости (рис. 1-10, а).

Пусть в качестве примеси в кристаллическую решетку полупроводника с ковалентной связью внесены элементы третьей группы периодической системы Менделеева, например алюминий в решетке кремния. Поскольку высшая валентность алюминия равна трем: $Al^{(13)} (1s^2 2s^2 2p^6 3s^2 3p)$, то одна связь атома кремния будет не завершена (рис. 1-9, б).

В незаполненную связь около атома алюминия за счет тепловой энергии 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 кристалла. *Примесь, захватывающая электроны, называется акцепторной.* Для образования свободной дырки за счет перехода электрона от атома основного вещества к атому примеси требуется значительно меньше энергии, чем для разрыва валентных связей кремния. В силу этого количество дырок может быть значительно больше количества свободных электронов, и проводимость такого кристалла будет дырочная. В таком полупроводнике основными носителями заряда будут

дырки, а электроны — неосновными носителями заряда. Полупроводник с акцепторными примесями 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 происходит переход электрона из валентной зоны на уровень E_a , а в валентной зоне появляется вакантная связь — дырка, которая и является свободным носителем заряда.

В полупроводниках могут одновременно содержаться как донорные, так и акцепторные примеси. Такие полупроводники называются компенсированными.

1-3. ПРОВОДИМОСТЬ ПОЛУПРОВОДНИКА

Проведем подсчет плотности тока для электронного полупроводника, электроны проводимости которого будем рассматривать как классические частицы, имеющие в каждый момент времени определенные положения и скорости.

Плотность тока есть заряд, переносимый в единицу времени через единицу поперечного сечения. Поэтому плотность тока j определяется концентрацией электронов проводимости n (число свободных электронов в единице объема вещества), средней скоростью их дрейфа $\langle v \rangle$ и величиной заряда e , т. е.

$$j = en \langle v \rangle. \quad (1-4)$$

Определение дрейфовой скорости проведем с учетом того, что тепловое равновесие электронов есть понятие статистическое, а поэтому и время свободного пробега должно иметь вероятностный характер.

Предположим, что dt/τ есть вероятность того, что электрон за время dt испытает столкновение (рассеяние). Кроме того, будем считать, что вероятность рассеяния электрона не зависит от времени, прошедшего с момента его последнего столкновения. Это значит, что τ есть некоторая постоянная величина. Тогда вероятность столкновений n частиц за время dt соответственно будет равна $n dt/\tau$. Следовательно, за время dt число носителей, движущихся в данном направлении, уменьшается в результате рассеяния на величину

$$-dn = n \frac{dt}{\tau}. \quad (1-5)$$

Решая уравнение (1-5) относительно n , получаем количество электронов, не испытавших к моменту времени t соударений:

$$n(t) = n_0 e^{-\frac{t}{\tau}}, \quad (1-6)$$

где $n = n_0$ при $t = 0$.

Из соотношения (1-6) следует, что количество электронов, движущихся в данном направлении, в результате столкновений уменьшается по экспоненциальному закону с постоянной времени τ .

Для определения τ поступим следующим образом. Предположим, что все электроны в момент времени $t = 0$ испытали столкновения. После этого в направлении x до следующего столкновения в результате дрейфа каждый носитель заряда прошел расстояние $x_1, x_2, x_3 \dots$ соответственно в течение времени свободного пробега $t_1, t_2, t_3 \dots$. Тогда носители заряда за время $T = t_1 + t_2 + t_3 + \dots$ пройдут суммарное расстояние, равное $X = x_1 + x_2 + x_3 + \dots$. Средняя скорость дрейфа носителей при этом составит:

$$\langle v \rangle = \frac{X}{T}. \quad (1-7)$$

Если n_0 электронов имеют среднее время свободного пробега $\langle t \rangle$, то

$$T = n_0 \langle t \rangle. \quad (1-8)$$

С другой стороны, есть определенная вероятность того, что среди электронов имеются такие, которые обладают одним и тем же временем свободного пробега t . Это будут электроны, испытавшие соударения в момент времени от t до $t + dt$. Их число согласно (1-5) составит $n \times dt / \tau$, и в суммарное время T они внесут часть $t \, ndt / \tau$. Тогда, интегрируя это выражение по всем временам свободного пробега, будем иметь:

$$T = \int_0^{\infty} \frac{tn \, dt}{\tau}. \quad (1-9)$$

Воспользовавшись выражениями (1-8), (1-9) и (1-6), получим:

$$\langle t \rangle = \frac{1}{n_0} \int_0^{\infty} t \frac{n \, dt}{\tau} = \frac{1}{n_0} \int_0^{\infty} \tau n_0 \frac{t}{\tau} e^{-\frac{t}{\tau}} \frac{dt}{\tau} = \tau. \quad (1-10)$$

Таким образом, величина τ имеет смысл среднего времени свободного пробега носителей заряда. За время свободного пробега t электрон, двигаясь с ускорением в электрическом поле с напряженностью \mathcal{E} , приобретет дрейфовую скорость, равную

$$v = \frac{e\mathcal{E}}{m} t, \quad (1-11)$$

и пройдет путь

$$x = \frac{e\mathcal{E}}{2m} t^2. \quad (1-12)$$

Суммарное расстояние, которое пройдут все электроны по направлению x , совпадающему с направлением внешнего поля,

будет:

$$X = x_1 + x_2 + \dots = \int_0^{\infty} xn \frac{dt}{\tau} = \frac{en_0\tau^2\mathcal{E}}{2m} \int_0^{\infty} \left(\frac{t}{\tau}\right)^2 e^{-\frac{t}{\tau}} \frac{dt}{\tau} = \frac{en_0\tau^2\mathcal{E}}{m}. \quad (1-13)$$

Следовательно, средняя скорость дрейфового движения равна:

$$\langle v \rangle = \frac{X}{T} = \frac{e\tau}{m} \mathcal{E}, \quad (1-14)$$

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

Величина скорости дрейфа в электрическом поле единичной напряженности, называемая подвижностью, определяется соотношением

$$\mu = \frac{\langle v \rangle}{\mathcal{E}} = \frac{e\tau}{m}. \quad (1-15)$$

С учетом соотношения (1-15) выражение (1-4) для плотности тока примет вид:

$$\mathbf{j} = en \langle \mathbf{v} \rangle = en\mu \mathcal{E}. \quad (1-16)$$

На основании закона Ома удельная электрическая проводимость может быть выражена с учетом (1-16) как

$$\sigma = j / \mathcal{E} = en\mu \quad (1-17)$$

или

$$\sigma = \frac{e^2 n \tau}{m}. \quad (1-18)$$

1.4. ЭФФЕКТ ХОЛЛА

Из формулы (1-17) следует, что измерение удельной проводимости позволяет вычислить произведение концентрации носителей заряда на их подвижность. Раздельно вычислить оба параметра полупроводника, измерив только электропроводность, не представляется возможным. Для решения этой задачи привлекаются данные, полученные при измерении эффекта Холла, сущность которого заключается в следующем.

Пусть по полупроводнику, имеющему форму прямоугольной пластинки, протекает электрический ток плотностью \mathbf{j} (рис. 1-11). В этом случае вектор плотности тока совпадает по направлению с вектором напряженности

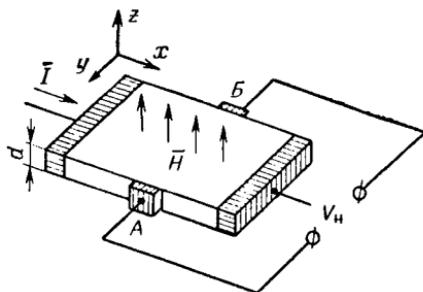


Рис. 1-11. Образец для измерения эффекта Холла.

приложенного к образцу электрического поля \mathcal{E} . Если полупроводник однороден, то эквипотенциальная поверхность, проходящая через AB на рис. 1-11, расположена перпендикулярно направлению электрического поля \mathcal{E} и, следовательно, вектору плотности тока \mathbf{j} . Поэтому разность потенциалов между точками A и B будет равна нулю.

Поместим теперь полупроводник в магнитное поле, перпендикулярное направлению тока, как это показано на рис. 1-11.

Как известно, на электрический заряд e , движущийся со скоростью \mathbf{v} в магнитном поле \mathbf{H} , действует сила Лоренца

$$\mathbf{F} = e\kappa\kappa_0 [\mathbf{v}\mathbf{H}] = e [\mathbf{V}\mathbf{B}], \quad (1-19)$$

где $\mathbf{B} = \kappa\kappa_0\mathbf{H}$ — магнитная индукция; κ — относительная магнитная проницаемость вещества; κ_0 — магнитная постоянная.

Направление этой силы зависит от знака носителей заряда e и векторного произведения скорости \mathbf{v} на магнитное поле \mathbf{H} . Если

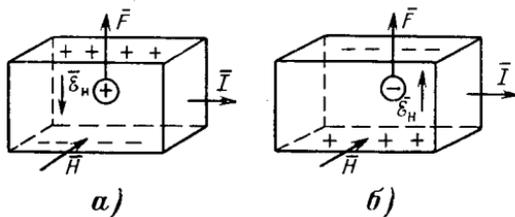


Рис. 1-12. Отклонение носителей заряда под воздействием магнитного поля в образцах с дырочной (а) и электронной (б) проводимостью.

скорость носителей заряда перпендикулярна магнитному полю \mathbf{H} , то под действием силы Лоренца произойдет отклонение носителей заряда в направлении, перпендикулярном \mathbf{v} и \mathbf{H} . В результате этого возникнет пространственное разделение зарядов и появится электрическое поле (рис. 1-12). В этом случае напряженность электрического поля \mathcal{E} уже не будет совпадать с направлением плотности тока \mathbf{j} , а следовательно, эквипотенциальные поверхности теперь не будут проходить через AB (рис. 1-11), и между точками A и B возникнет разность потенциалов. Это явление называется эффектом Холла, а поперечная разность потенциалов — напряжением Холла.

Под действием силы Лоренца для выбранного нами направления \mathbf{H} и \mathcal{E} дырки в акцепторном полупроводнике (рис. 1-12, а) и электроны в полупроводнике (рис. 1-12, б) будут отклоняться к верхней грани образца, а на нижней грани будет недостаток соответствующих носителей заряда, что обусловит противоположный по знаку заряд по отношению заряда на верхней грани. Этот процесс будет продолжаться до тех пор, пока возникшее за счет разделения носителей заряда поперечное электрическое поле \mathcal{E}_H не создаст силы, действующей на свободные носители заряда и уравновешивающей силу Лоренца. В стационарном состоянии эти силы равны и для электронного полупроводника запишутся в виде

$$-e\mathcal{E}_H = e\mathbf{v}\mathbf{B}. \quad (1-20)$$

Если ширина образца d , то холловская разность потенциалов V_H между противоположными гранями образца составит величину

$$V_H = \mathcal{E}_H d = -vBd. \quad (1-21)$$

Выразив величину v согласно соотношению (1-4) через плотность тока j , будем иметь:

$$V_H = -\frac{1}{en} jBd = RjBd. \quad (1-22)$$

Величина R в уравнении (1-22) называется коэффициентом Холла, и в случае электронов она равна:

$$R = -\frac{1}{en}. \quad (1-23)$$

Если носителями заряда являются дырки с концентрацией p , то, как видно из рис. 1-12, а, $e\mathcal{E}_H = evB$, поэтому

$$V_H = \frac{1}{ep} jBd = RjBd \quad (1-24)$$

и

$$R = \frac{1}{ep}. \quad (1-25)$$

Таким образом, зная коэффициент Холла, можно вычислить концентрацию носителей заряда и определить их знак.

Пусть полупроводниковый образец ориентирован в пространстве так, что вектор электрического поля \mathcal{E} совпадает с направлением оси x , а вектор напряженности магнитного поля \vec{H} с направлением оси z , как это изображено на рис. 1-13. Рассмотрим случай слабых магнитных полей, когда угол φ между векторами \vec{j} и \mathcal{E} , называемый углом Холла, мал.

Разложим вектор напряженности электрического поля \mathcal{E} на две перпендикулярные составляющие \mathcal{E}_t и \mathcal{E}_l , как это показано на рис. 1-13. При этом, учитывая малость угла φ , имеем:

$$\mathcal{E}_l = \mathcal{E}_t \operatorname{tg} \varphi = \mathcal{E}_t \varphi. \quad (1-26)$$

Если проводимость образца σ , то $\mathcal{E}_l = j/\sigma$. Тогда холловская разность потенциалов будет равна:

$$V_H = \mathcal{E}_l d = \frac{j}{\sigma} \varphi d = \frac{\varphi}{\sigma B} jBd, \quad (1-27)$$

откуда следует, что

$$R = \frac{\varphi}{\sigma B}. \quad (1-28)$$

$$D = \epsilon_0 \epsilon \frac{E_0}{\kappa \rho}$$

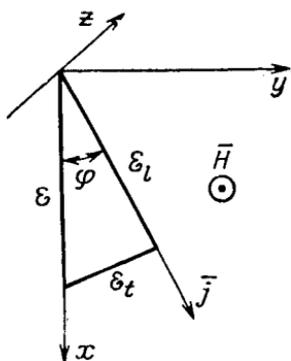


Рис. 1-13. К выводу выражения для коэффициента Холла.

Если учесть равенства (1-17) и (1-23), то угол Холла определится выражением вида

$$\varphi = \mu_n B, \quad (1-29)$$

где μ_n — холловская подвижность носителей заряда.

Зная для одного и того же образца коэффициент Холла и удельную проводимость, согласно (1-28) и (1-29) можно вычислить холловскую подвижность носителей заряда

$$\mu_n = |R| \sigma. \quad (1-30)$$

В таком рассмотрении эффекта Холла не принималось во внимание статистическое распределение носителей заряда по скоростям. С учетом механизма рассеяния носителей заряда выражение для коэффициента Холла запишется в виде

$$|R| = r \frac{1}{en}. \quad (1-31)$$

Для атомных полупроводников при рассеянии носителей заряда на акустических колебаниях кристаллической решетки коэффициент пропорциональности

$$r = \frac{3\pi}{8},$$

в случае же рассеяния на ионах примеси он принимает значение

$$r = 1,93,$$

а для рассеяния на нейтральной примеси

$$r = 1.$$

Проводимость полупроводника согласно (1-17) определяется дрейфовой подвижностью μ_d . В то же время, если имеется, кроме того, еще и магнитное поле, движение носителей заряда характеризуется холловской подвижностью μ_n . При этом холловская подвижность связана с дрейфовой согласно равенствам (1-17), (1-30) и (1-31) соотношением вида

$$\mu_n = r \mu_d. \quad (1-32)$$

1-5. ЭФФЕКТ ХОЛЛА ДЛЯ ПОЛУПРОВОДНИКА СО СМЕШАННОЙ ПРОВОДИМОСТЬЮ

Пусть в полупроводнике ток переносится отрицательными и положительными носителями заряда, концентрации которых соответственно равны n и p . Такой полупроводник называется полупроводником со смешанной проводимостью.

Определение коэффициента Холла проведем для случая слабого магнитного поля. Кроме того, будем считать, что движение носителей заряда определяется только приложенными внешними силовыми полями, т. е. после каждого соударения носители заряда

приходят в состояние покоя. Эти допущения позволяют воспользоваться полученными закономерностями в эффекте Холла для носителей заряда одного знака. Пусть вектор напряженности электрического поля \mathcal{E} , вектор плотности тока \mathbf{j} и вектор напряженности магнитного поля \mathbf{H} ориентированы относительно осей координат, как это показано на рис. 1-14.

Под действием электрического поля носители заряда будут ускоряться. По мере увеличения скорости на них будет действовать отклоняющая сила Лоренца. Траектория носителя заряда искривляется до тех пор, пока он не испытает столкновения с неоднородностями кристаллической решетки, после чего процесс начи-

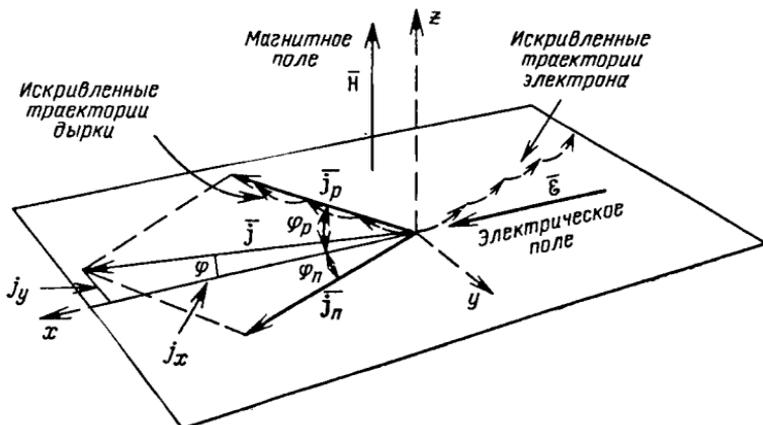


Рис. 1-14. Эффект Холла при наличии двух типов носителей заряда.

нается сначала. В результате ток дырок отклоняется от направления электрического поля на угол φ_p . Электроны под действием магнитного поля отклоняются в ту же сторону, что и дырки, а электронный ток отклоняется от направления электрического поля в противоположном направлении на угол φ_n .

В силу этого движение носителей заряда обусловит появление токов плотности $\mathbf{j}_p = \sigma_p \mathcal{E}$ и $\mathbf{j}_n = \sigma_n \mathcal{E}$, которые будут направлены, как показано на рис. 1-14. При этом суммарная плотность тока будет равна:

$$\mathbf{j} = \mathbf{j}_n + \mathbf{j}_p. \quad (1-33)$$

Поскольку магнитное поле слабое, то углы Холла φ_n для электронов и φ_p для дырок малы. Поэтому на основании выражения (1-29) они могут быть записаны в виде

$$\varphi_p = -\mu_p B, \quad (1-34)$$

$$\varphi_n = \mu_n B. \quad (1-35)$$

Суммарный угол Холла, исходя из рис. 1-14, равен:

$$\operatorname{tg} \varphi \approx \varphi = \frac{-j_y}{j_x}. \quad (1-36)$$

Из рис. 1-14 следует, что

$$j_x = j_p \cos \varphi_p + j_n \cos \varphi_n = e (\rho \mu_p + n \mu_n) \mathcal{E} = \sigma \mathcal{E}, \quad (1-37)$$

где $\cos \varphi_p = \cos \varphi_n = 1$ из-за малости углов; $\sigma_p = e \rho \mu_p$, $\sigma_n = e n \mu_n$ — проводимость, обусловленная дырками и электронами соответственно; $\sigma = \sigma_n + \sigma_p$ — суммарная проводимость;

$$j_y = j_p \sin \varphi_p + j_n \sin \varphi_n = j_p \varphi_p + j_n \varphi_n = -(\sigma_p \mu_p - \sigma_n \mu_n) \mathcal{E} B = \\ = -e (\rho \mu_p^2 - n \mu_n^2) \mathcal{E} B, \quad (1-38)$$

где $\sin \varphi_p \approx \varphi_p$ и $\sin \varphi_n \approx \varphi_n$ из-за малости напряженности магнитного поля.

Таким образом, в случае смешанной проводимости угол Холла

$$\varphi = \frac{-j_y}{i_x} = \frac{\rho \mu_p^2 - n \mu_n^2}{\rho \mu_p + n \mu_n} B, \quad (1-39)$$

а коэффициент Холла

$$R = \frac{\varphi}{\sigma B} = \frac{\rho \mu_p^2 - n \mu_n^2}{(\rho \mu_p + n \mu_n)^2} \cdot \frac{1}{e}. \quad (1-40)$$

В случае смешанной проводимости, как следует из соотношения (1-40), знак коэффициента Холла определяется соотношением концентраций и подвижностей электронов и дырок. Холловская подвижность такого полупроводника равна:

$$\mu_H = R \sigma = \frac{\rho \mu_p^2 - n \mu_n^2}{\rho \mu_p + n \mu_n}. \quad (1-41)$$

В собственном полупроводнике концентрации свободных электронов и дырок равны, т. е. $n = p = n_i$, поэтому коэффициент Холла

$$R_i = \frac{\mu_p - \mu_n}{n_i (\mu_p + \mu_n)} \cdot \frac{1}{e}. \quad (1-42)$$

Чаще всего он по знаку отрицателен из-за большей подвижности электронов.

Глава вторая

ОСНОВЫ ЗОННОЙ ТЕОРИИ ПОЛУПРОВОДНИКОВ

2-1. УРАВНЕНИЕ ШРЕДИНГЕРА ДЛЯ КРИСТАЛЛА

Любое твердое тело представляет собой совокупность большого числа атомных ядер и электронов. Для определения стационарных состояний и энергетического спектра такой системы частиц нужно решить уравнение Шредингера для кристалла

$$\hat{H}\Psi = E\Psi, \quad (2-1)$$

где E — энергия кристалла; Ψ — его собственная волновая функция, зависящая от координат всех электронов \mathbf{r}_i и атомных ядер \mathbf{R}_α ,

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

\hat{H} — оператор Гамильтона, включающий в себя следующие виды энергии:

1) кинетическую энергию электронов

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

где \hbar — постоянная Планка h , деленная на 2π ; m — 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) кинетическую энергию ядер

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

где M_α — масса ядра;

3) потенциальную энергию попарного взаимодействия электронов между собой

$$\frac{1}{2} \sum_i \sum_{\substack{j \\ i \neq j}} \frac{e^2}{4\pi\epsilon\epsilon_0 r_{ij}},$$

где ϵ — диэлектрическая проницаемость кристалла; ϵ_0 — диэлектрическая постоянная;

4) потенциальную энергию попарного взаимодействия ядер между собой

$$\mathcal{V}_0(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N);$$

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:

$$\left\{ \sum_i \left(-\frac{\hbar^2}{2m} \Delta_i \right) + \sum_\alpha \left(-\frac{\hbar^2}{2M_\alpha} \Delta_\alpha \right) + \frac{1}{2} \sum_i \sum_{\substack{j \\ i \neq j}} \frac{e^2}{4\pi\epsilon\epsilon_0 r_{ij}} + \mathcal{V}_0(\mathbf{R}_1, \mathbf{R}_2, \dots) + \right. \\ \left. + U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n; \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \right\} \Psi = E\Psi. \quad (2-2)$$

Число независимых переменных в этом уравнении определяется полным числом частиц в кристалле (порядка 10^{23} см^{-3}). Поэтому непосредственное решение уравнения (2-2) встречает математиче-

ские трудности. Возможно лишь приближенное решение задачи, 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 тех и других частиц одного порядка. Так как масса ядра намного больше массы электрона, т. е. $M \gg m$, 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. Это позволяет в первом приближении рассматривать движение электронов в потенциальном поле фиксированных ядер. При изучении движения ядер, напротив, следует учитывать не мгновенное положение электронов, а поле, создаваемое их средним пространственным распределением. Такое приближенное рассмотрение называется адиабатическим или приближением Борна-Оппенгеймера.

Предположим, что ядра покоятся. Для этого случая координаты ядер R_1, R_2, \dots, R_N уже не переменные, а параметры $R_{10}, R_{20}, \dots, R_{N0}$, представляющие собой координаты узлов кристаллической решетки. Для случая покоящихся ядер, когда $R_\alpha = R_{\alpha 0}$, кинетическая энергия ядер обращается в нуль, а энергия их взаимодействия \mathcal{V}_0 принимает постоянное значение. Выбором начала отсчета энергии \mathcal{V}_0 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 и их волновая функция Ψ_e параметрически зависят от расположения ядер $R_{\alpha 0}$. В результате уравнение Шредингера для электронов запишется в виде

$$\left\{ \sum_i \left(-\frac{\hbar^2}{2m} \Delta_i \right) + \frac{1}{2} \sum_i \sum_{j, i \neq j} \frac{e^2}{4\pi\epsilon_0 r_{ij}} + U(r_1, r_2, \dots, r_n; R_{10}, R_{20}, \dots, R_{N0}) \right\} \Psi_e = E \Psi_e. \quad (2-3)$$

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 уравнение (2-3) записывается только для валентных электронов, движущихся в потенциальном поле фиксированных ядер.

Несмотря на проведенные упрощения, уравнение (2-3) не может быть решено в общем виде. По-прежнему имеем дело с задачей многих частиц, но 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 этом поле будет Ω_i . Она зависит не только от движения всех остальных электронов, но и от движения данного электрона, так как его движение оказывает влияние на движение остальных электронов. Поскольку поле Ω_i не только определяет движение данного электрона, но и зависит от его движения, то это поле получило название с а м о с о г л а с о в а н н о г о. Поле Ω_i может быть определено методом последовательных приближений. Введение самосогласованного поля позволяет в уравнении (2-3) двойную сумму $\frac{1}{2} \sum_i \sum_{j \neq i} \frac{e^2}{4\pi\epsilon_0 r_{ij}}$ заменить суммой $\sum_i \Omega_i(\mathbf{r}_i)$, каждый

член которой зависит от координат одного электрона. В этом случае $\Omega_i(\mathbf{r}_i)$ является потенциальной энергией i -го электрона в поле всех остальных электронов.

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

$$U(\mathbf{r}_1, \mathbf{r}_2 \dots) = \sum_i U_i(\mathbf{r}_i), \quad (2-4)$$

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

С учетом (2-4) уравнение (2-3) преобразуется к виду

$$\left\{ \sum_i \left[-\frac{\hbar^2}{2m} \Delta_i + \Omega_i(\mathbf{r}_i) + U_i(\mathbf{r}_i) \right] \right\} \Psi_e = E \Psi_e. \quad (2-5)$$

Это уравнение можно записать как

$$\left(\sum_i \hat{H}_i \right) \Psi_e = E \Psi_e, \quad (2-6)$$

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

$$\hat{H}_i = -\frac{\hbar^2}{2m} \Delta_i + \Omega_i(\mathbf{r}_i) + U_i(\mathbf{r}_i). \quad (2-7)$$

Из равенства (2-6) следует, что электроны движутся независимо друг от друга в некотором самосогласованном потенциальном поле. Таким образом, введение самосогласованного поля позволяет элек-

троны кристалла рассматривать как систему взаимодействующих частиц. Поскольку гамильтониан не содержит энергии взаимодействия электронов, то волновую функцию системы электронов можно представить в виде произведения волновых функций отдельных электронов:

$$\Psi_e(\mathbf{r}_1, \mathbf{r}_2 \dots) = \Psi_1(\mathbf{r}_1) \Psi_2(\mathbf{r}_2) \dots = \Pi \Psi_i, \quad (2-8)$$

а полная энергия системы равна сумме энергий отдельных электронов:

$$E = E_1 + E_2 + \dots = \sum_i E_i. \quad (2-9)$$

Обозначим через $\Pi_{(1)} \Psi_i$ произведение всех волновых функций, кроме Ψ_1 , т. е.

$$[\Pi_{(1)} \Psi_i] \Psi_1 = [\Pi_{(2)} \Psi_i] \Psi_2 = \dots = \Pi \Psi_i. \quad (2-10)$$

Учитывая, что оператор \hat{H}_i действует только на волновую функцию Ψ_i , из уравнений (2-6), (2-8) получаем:

$$[\Pi_{(1)} \Psi_i] \hat{H}_1 \Psi_1 + [\Pi_{(2)} \Psi_i] \hat{H}_2 \Psi_2 + \dots = E \Pi \Psi_i. \quad (2-11)$$

Разделим обе части уравнения (2-11) на $\Pi \Psi_i$:

$$\frac{1}{\Psi_1} \hat{H}_1 \Psi_1 + \frac{1}{\Psi_2} \hat{H}_2 \Psi_2 + \dots = E_1 + E_2 + \dots \quad (2-12)$$

Каждый член в левой части уравнения (2-12) зависит от координат лишь одного электрона. Это уравнение эквивалентно системе следующих одноэлектронных уравнений:

$$\begin{aligned} \hat{H}_1 \Psi_1 &= E_1 \Psi_1; \\ \hat{H}_2 \Psi_2 &= E_2 \Psi_2. \\ &\vdots \\ &\vdots \end{aligned} \quad (2-13)$$

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

$$\hat{H} \Psi(\mathbf{r}) = E \Psi(\mathbf{r}), \quad (2-14)$$

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

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

$$U(\mathbf{r}) + \Omega(\mathbf{r}) = V(\mathbf{r}), \quad (2-15)$$

то уравнение Шредингера для электрона кристалла запишется в виде

$$\left\{ -\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}) \right\} \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2-16)$$

Существуют различные способы рационального выбора вида функции $\Psi(\mathbf{r})$ 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 (2-15) и (2-16) есть некая функция, обладающая трехмерной периодичностью кристаллической решетки. Качественный вид ее для одномерной решетки показан на рис. 2-1, где пунктирной кривой изображен потенциал изолированного атома $U_a(\mathbf{r})$.

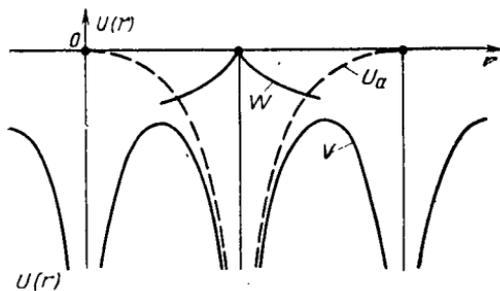


Рис. 2-1. Схематическое изображение периодического потенциала одномерной решетки.

Блох показал, что для случая периодического потенциального поля решение уравнения Шредингера имеет вид:

$$\Psi_{\mathbf{k}}(\mathbf{r}) = U_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}}, \quad (2-17)$$

где $U_{\mathbf{k}}(\mathbf{r})$ обладает трехмерной периодичностью решетки кристалла; \mathbf{k} — волновой вектор электрона в кристалле.

Его размерность — обратная длина. Член $U_{\mathbf{k}}(\mathbf{r})$ определяется конкретным видом функции $V(\mathbf{r})$, но оказывается, что независимо от этого из одного условия периодичности $U_{\mathbf{k}}(\mathbf{r})$ можно сделать важные выводы о состояниях электронов в кристалле.

2-4. ПРИБЛИЖЕНИЕ СИЛЬНО СВЯЗАННЫХ ЭЛЕКТРОНОВ

Если считать, что кинетическая энергия электрона значительно больше пространственных изменений потенциальной энергии $V(\mathbf{r})$, то влияние периодического потенциала можно рассматривать как малое возмущение свободного движения электронов. Такой подход, получивший название «приближение 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 хорош только для электронов глубоких энергетических уровней. Ни одно из этих приближений не позволяет количественно описать поведение валентных электронов, но они хорошо иллюстрируют общие закономерности движения электрона в периодическом поле кристалла.

Пусть уравнение Шредингера для изолированного атома имеет вид:

$$\hat{H}_a \Psi_a = E_a \Psi_a, \quad (2-18)$$

где E_a — энергия некоторого уровня; Ψ_a — соответствующая уровню E_a волновая функция.

Будем искать решение уравнения (2-16) в приближении сильно связанных электронов в виде линейной комбинации атомных волновых функций

$$\Psi(\mathbf{r}) = \sum_g a_g \Psi_g, \quad (2-19)$$

где

$$\Psi_g = \Psi_a(\mathbf{r} - \mathbf{R}_g), \quad (2-20)$$

\mathbf{R}_g — радиус-вектор узла кристаллической решетки.

Чтобы выражение (2-19) для $\Psi(\mathbf{r})$ удовлетворяло виду блоховской функции (2-17), необходимо положить:

$$a_g = C e^{i\mathbf{k}\mathbf{R}_g}, \quad (2-21)$$

где C — постоянная величина.

Подставив в уравнение (2-14) выражение (2-19), получим:

$$\sum_g a_g \{ \hat{H} \Psi_g - E \Psi_g \} = 0. \quad (2-22)$$

Гамильтониан для изолированного атома имеет вид:

$$\hat{H}_a = -\frac{\hbar^2}{2m} \Delta + U_a(\mathbf{r}), \quad (2-23)$$

где $U_a(\mathbf{r})$ — значение потенциальной энергии электрона в атоме. Для кристалла

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}), \quad (2-24)$$

где $V(\mathbf{r})$ — периодическое поле кристаллической решетки.

Вычитая выражение (2-23) из (2-24), будем иметь:

$$\hat{H} = \hat{H}_a + W(\mathbf{r}). \quad (2-25)$$

Здесь через

$$W(\mathbf{r}) = V(\mathbf{r}) - U_a(\mathbf{r}) \quad (2-26)$$

обозначена энергия возмущения для электрона в кристалле по сравнению с изолированным атомом (рис. 2-1).

Уравнение Шредингера (2-22) с учетом (2-25) примет вид:

$$\sum_g a_g \{ \hat{H}_a \Psi_g + W(\mathbf{r}) \Psi_g - E \Psi_g \} = 0. \quad (2-27)$$

Если учесть выражения (2-18) и (2-20), то это равенство преобразуется следующим образом:

$$\sum_g a_g \{W(\mathbf{r}) \Psi_g + (E_a - E) \Psi_g\} = 0. \quad (2-28)$$

Поставим задачу найти собственные значения энергии электрона E в кристалле, считая, что энергетический спектр изолированного атома E_a известен. Умножим уравнение (2-28) на функцию $\Psi_{g'}^*$, комплексно сопряженную с Ψ_g , и проинтегрируем по всему объему кристалла τ :

$$\sum_g a_g \left\{ \int_{\tau} \Psi_g W(\mathbf{r}) \Psi_{g'}^* d\tau + (E_a - E) \int_{\tau} \Psi_g \Psi_{g'}^* d\tau \right\} = 0. \quad (2-29)$$

Введем обозначения

$$\int_{\tau} \Psi_g W(\mathbf{r}) \Psi_{g'}^* d\tau = A(\mathbf{R}_g - \mathbf{R}_{g'}). \quad (2-30)$$

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

$$\int_{\tau} \Psi_g \Psi_{g'}^* d\tau = S(\mathbf{R}_g - \mathbf{R}_{g'}). \quad (2-31)$$

Этот интеграл зависит только от перекрытия атомных волновых функций.

С учетом уравнения (2-31) соотношение (2-29) можно записать в виде

$$\sum_g A(\mathbf{R}_g - \mathbf{R}_{g'}) e^{i\mathbf{k}\mathbf{R}_g} + (E_a - E) \sum_g S(\mathbf{R}_g - \mathbf{R}_{g'}) e^{i\mathbf{k}\mathbf{R}_g} = 0. \quad (2-32)$$

Умножим полученное выражение на $e^{-i\mathbf{k}\mathbf{R}_{g'}}$ [заметим, что суммирование в уравнении (2-32) производится только по g , а g' — постоянная величина]. Тогда

$$\sum_g A(\mathbf{R}_g - \mathbf{R}_{g'}) e^{i\mathbf{k}(\mathbf{R}_g - \mathbf{R}_{g'})} + (E_a - E) \sum_g S(\mathbf{R}_g - \mathbf{R}_{g'}) e^{i\mathbf{k}(\mathbf{R}_g - \mathbf{R}_{g'})} = 0. \quad (2-33)$$

Упростим уравнение (2-33), введя вектор \mathbf{q} , соединяющий атомы g и g' ,

$$\mathbf{R}_g - \mathbf{R}_{g'} = \mathbf{q}. \quad (2-34)$$

При этом получим:

$$\sum_g A(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}} + (E_a - E) \sum_g S(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}} = 0. \quad (2-35)$$

Отсюда можно выразить значение энергии электрона в кристалле

$$E = E_a + \frac{\sum_g A(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}}}{\sum_g S(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}}}. \quad (2-36)$$

Из равенства (2-36) 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 k . Теперь вместо одного атомного уровня в кристалле имеем энергетическую зону, границы которой определяются минимальным и максимальным значением второго члена слагаемого.

Рассмотрим более подробно добавочный член в выражении (2-36) для случая простой одномерной решетки с периодом a (рис. 2-2).

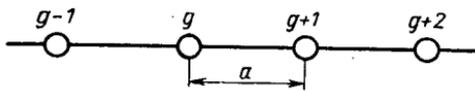


Рис. 2-2. Схематическое изображение одномерной решетки.

Для упрощения задачи будем считать, что благодаря быстрому спаданию атомных волновых функций можно

пренебречь их перекрытием даже для соседних атомов. Такое допущение является обоснованным в приближении сильно связанных электронов. При этом все члены в выражении $S(q)$ обращаются в нуль, кроме члена с $q = 0$, который равен единице из условия нормировки атомных волновых функций, т. е.

$$S(q) = \begin{cases} 1 & q = 0, \\ 0 & q \neq 0, \end{cases} \quad (2-37)$$

откуда следует, что

$$\sum_g S(q) = 1. \quad (2-38)$$

В сумме, стоящей в числителе дроби (2-36), следует также учитывать в первом приближении слагаемые для соседних узлов ($q = a$ и $q = -a$). Действительно, малое значение волновой функции атома вблизи соседнего узла частично компенсируется большой разностью $W(r) = V(r) - U_a(r)$ в этой области (см. рис. 2-1). Поэтому числитель дроби можно записать следующим образом:

$$\sum_g A(q) e^{ikq} = A(0) + A(+a) e^{ika} + A(-a) e^{-ika}. \quad (2-39)$$

Первый член суммы

$$A(0) = \int_{\tau} \Psi_0(r) W(r) \Psi_0^*(r) d\tau = \int_{\tau} |\Psi_0(r)|^2 W(r) d\tau = C \quad (2-40)$$

есть постоянная величина, не зависящая от k . Значение обменных интегралов $A(+a)$ и $A(-a)$ зависит от степени перекрытия электронных облаков соседних атомов. Для идеальной решетки эти члены одинаковы из условия идентичности всех узлов, т. е.

$$A(+a) = A(-a) = A. \quad (2-41)$$

Используя введенные обозначения, имеем:

$$\sum_g A(q) e^{ikq} = C + A e^{ika} + A e^{-ika} = C + 2A \cos ka. \quad (2-42)$$

Выражение (2-36) для энергии электрона в периодическом поле одномерной решетки принимает вид:

$$E = E_a + C + 2A \cos ka. \quad (2-43)$$

Обобщая это равенство для случая трехмерной кубической решетки, получаем:

$$E = E_a + C + 2A (\cos k_x a + \cos k_y a + \cos k_z a). \quad (2-44)$$

Анализ данного выражения позволяет 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 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 (2-44), имеющие место при $\cos k_i a = \pm 1$ ($i = x, y, z$), будут:

$$E_{\text{макс}} = E_a + C + 6A;$$

$$E_{\text{мин}} = E_a + C - 6A.$$

Следовательно, для простой кубической решетки ширина энергетической зоны равна:

$$E_{\text{макс}} - E_{\text{мин}} = 12A, \quad (2-45)$$

т. е. ширина энергетической зоны зависит от величины обменного интеграла.

4. Если в изолированном атоме имеется несколько уровней, то в кристалле каждый из них расщепляется в зону. Поскольку величина обменного интеграла определяется перекрытием электронных облаков соседних атомов, то, чем сильнее перекрываются волновые функции атомов, тем больше величина A и, следовательно, ширина энергетической зоны. В силу этого для более высоких уровней из-за большего перекрытия волновых функций образуется более широкая энергетическая зона (рис. 2-3).

5. Энергетические зоны в общем случае разделены интервалами энергии ΔE_g , называемыми запрещенными зонами (рис. 2-3).

6. С ростом энергии ширина запрещенной зоны уменьшается, так как имеет место сильное расщепление высоко лежащих атомных уровней (рис. 2-3).

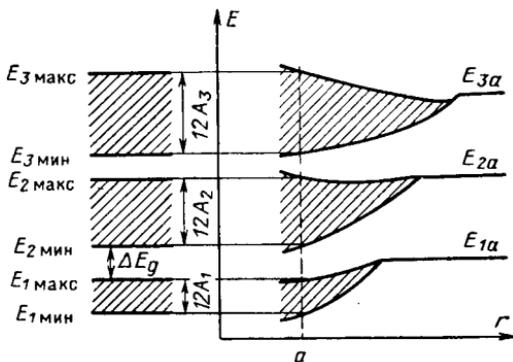


Рис. 2-3. Образование зон энергии из энергетических уровней при сближении атомов. a — постоянная решетки кристалла.

7. Уровень E_a в изолированном атоме может быть вырожден. В кристаллической решетке вырождение может быть частично или полностью снято. При этом атомный уровень расщепляется на несколько зон, число которых соответствует степени вырождения. Например, для p -состояния в кристалле возможно наличие трех зон.

8. Состояние электронов в кристалле зависит от компонент волнового вектора \mathbf{k} . Функция $E(\mathbf{k})$ периодическая и четная, т. е. $E(\mathbf{k}) = E(-\mathbf{k})$.

9. При воздействии на кристалл температуры и давления, приводящих к изменению расстояния между его атомами, будет изменяться область перекрытия волновых функций и, следовательно, величина обменного интеграла. Это вызовет изменение ширины зоны энергии и ширины запрещенной зоны.

2-5. ЧИСЛО СОСТОЯНИЙ В РАЗРЕШЕННОЙ ЗОНЕ

Для подсчета числа состояний в зоне нужно при решении уравнения Шредингера для электрона учесть граничные условия на краях кристалла. Любые условия на границах не отразятся существенным образом на состоянии электрона в объеме, так как число узлов в любом реальном кристалле велико. Наиболее удобным для решения поставленной задачи является применение условия цикличности Борна — Кармана.

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

$$L_x = aN_x; \quad L_y = aN_y; \quad L_z = aN_z, \quad (2-46)$$

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:

$$\Psi(x + L_x, y + L_y, z + L_z) = \Psi(x, y, z). \quad (2-47)$$

Уравнение (2-47) по существу и является циклическими граничными условиями Борна-Кармана. Учтя вид волновой функции для кристалла (2-17), получаем:

$$\begin{aligned} \Psi(x + L_x, y + L_y, z + L_z) &= U_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}} e^{i\mathbf{k}_x L_x} e^{i\mathbf{k}_y L_y} e^{i\mathbf{k}_z L_z} = \\ &= \Psi(x, y, z) e^{i\mathbf{k}_x L_x} e^{i\mathbf{k}_y L_y} e^{i\mathbf{k}_z L_z}. \end{aligned} \quad (2-48)$$

Чтобы выполнить условие (2-47), необходимо потребовать

$$e^{i\mathbf{k}_x L_x} = e^{i\mathbf{k}_y L_y} = e^{i\mathbf{k}_z L_z} = 1.$$

Это равенство выполнимо, если

$$\kappa_x L_x = 2\pi n_1; \quad \kappa_y L_y = 2\pi n_2; \quad \kappa_z L_z = 2\pi n_3,$$

где n_1, n_2, n_3 — любое целое число.

Отсюда следует, что

$$\kappa_x = \frac{2\pi}{L_x} n_1; \quad \kappa_y = \frac{2\pi}{L_y} n_2; \quad \kappa_z = \frac{2\pi}{L_z} n_3, \quad (2-49)$$

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

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

С учетом значений вектора \mathbf{k} (2-49) можно записать волновую функцию (2-19) для одномерной задачи в виде

$$\Psi_{\mathbf{k}} = C \sum_g e^{i\kappa_x a g} \Psi_g = C \sum_g e^{\frac{i2\pi n_1 g}{N_x}} \Psi_g, \quad (2-50)$$

где проведена замена $R_g = ag; g$ — целое число.

Из выражения (2-50) видно, что волновая функция для $n_1 = N_x$ совпадает с функцией для $n_1 = 0$, волновая функция для $n_1 = N_x + 1$ — с функцией для $n_1 = 1$ и т. д. Это означает, что компонента κ_x имеет N_x значений, соответствующих различным n_1 . При этом n_1 может принимать значения $n_1 = 0, 1, 2 \dots (N_x - 1)$, т. е. n_1 изменяется только в пределах.

$$0 \leq n_1 < N_x \quad \text{или} \quad -\frac{N_x}{2} \leq n_1 < +\frac{N_x}{2}. \quad (2-51)$$

Используя соотношения (2-51), (2-49) и (2-46), получаем пределы изменения для κ_x :

$$-\frac{\pi}{a} \leq \kappa_x < +\frac{\pi}{a}. \quad (2-52)$$

Аналогично для k_y и k_z можно написать:

$$\begin{aligned} -\frac{\pi}{a} &\leq \kappa_y < +\frac{\pi}{a}; \\ -\frac{\pi}{a} &\leq \kappa_z < +\frac{\pi}{a}, \end{aligned} \quad (2-53)$$

где k_y и k_z принимают соответственно N_y и N_z различных значений.

Следовательно, в разрешенной зоне трехмерного кристалла имеется всего $N = N_x N_y N_z$ энергетических состояний, соответствующих различным \mathbf{k} , т. е. равное числу атомов в кристалле. На каждом энергетическом уровне изолированного атома в соответствии с принципом Паули может находиться не более двух электронов с противоположно направленными спинами. Поэтому в каж-

дом энергетическом состоянии кристалла, которое определяется значением вектора \mathbf{k} , располагаются два электрона. Таким образом, в простой энергетической зоне имеется $2N$ квантовых состояний, и в ней может находиться $2N$ электронов. Если зона g кратно выражена, то в ней может быть $2Ng$ электронов. Значит в вырожденной зоне число квантовых состояний равно общему числу мест на уровнях изолированных атомов, из которых образовалась эта зона.

Аналогично соотношению де Бройля для свободного электрона для электрона, движущегося в периодическом поле кристалла, можно ввести величину $\mathbf{p} = \hbar\mathbf{k}$, называемую к в а з и м п у л ь с о м. В соответствии с дискретным спектром \mathbf{k} квазиимпульс также квантован и согласно неравенствам (2-52) и (2-53) изменяется в пределах

$$-\frac{\pi\hbar}{a} \leq p_i < +\frac{\pi\hbar}{a}; \quad i = x, y, z. \quad (2-54)$$

В отличие от выражения для энергии свободного электрона

$$E = \frac{p^2}{2m} = \frac{\hbar^2 \kappa^2}{2m} \quad (2-55)$$

для электрона в кристалле получим:

$$E = E_a + C + 2A \left[\cos\left(\frac{a}{\hbar} p_x\right) + \cos\left(\frac{a}{\hbar} p_y\right) + \cos\left(\frac{a}{\hbar} p_z\right) \right]. \quad (2-56)$$

Из соотношения (2-56) следует, что энергия электрона в кристалле — четная функция квазиимпульса, т. е.

$$E(+p) = E(-p). \quad (2-57)$$

Оценим среднее расстояние между соседними уровнями энергии в разрешенной зоне. Примем значение параметра решетки $a \approx 4\text{Å} = 4 \cdot 10^8 \text{ см}$, что дает объем элементарной ячейки $a_3 \approx 64 \cdot 10^{-24} \text{ см}^3$. Число элементарных ячеек в кристалле единичного объема $V = 1 \text{ см}^3$, равное числу состояний в зоне, составит:

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

При ширине зоны в 1 эв среднее расстояние между ее уровнями имеет порядок 10^{-22} эв, т. е. зону можно считать квазинепрерывной.

Средняя кинетическая энергия электрона в термодинамическом равновесии равна $3/2 k_0 T$ (k_0 — постоянная Больцмана) и при комнатной температуре ($T = 300^\circ \text{ К}$) составляет величину около 0,04 эв. Поэтому в результате теплового возбуждения электроны могут легко переходить на более высокие энергетические уровни, если они свободны.

Рассмотрим два частных случая, возникающих при образовании валентной зоны.

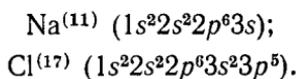
1. Число возможных состояний для электронов в зоне равно количеству валентных электронов атомов, образовавших кристалл.

При этом все уровни энергии в зоне заняты, и на каждом уровне находится по два электрона. Поэтому электроны в результате тепловых колебаний решетки могут лишь обмениваться местами, но их общая энергия не изменяется. При приложении электрического поля суммарная энергия электронов также не может измениться. Средняя скорость электронов остается равной нулю, так как на каждом уровне располагаются два электрона с противоположно направленными импульсами. Такой кристалл не может проводить ток, т. е. является диэлектриком.

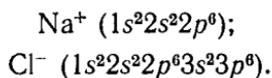
2. Зона заполнена электронами лишь частично или перекрывается со следующей свободной зоной. При температуре, отличной от абсолютного нуля, электроны в результате теплового возбуждения будут переходить на более высокие энергетические уровни. Но при отсутствии внешнего поля они движутся хаотически, и их средняя скорость равна нулю.

Если такой кристалл поместить в электрическое поле, то часть электронов, которые под действием тепла перешли на более высокие незанятые энергетические уровни, приобретет дополнительную скорость вдоль поля. В результате этого средняя скорость электронов становится отличной от нуля и возникает электрический ток. Это типичный механизм проводимости для металлов.

Разберем несколько конкретных примеров зонной структуры. Рассмотрим энергетический спектр хлористого натрия. Электронная структура изолированных атомов натрия и хлора записывается символически следующим образом:



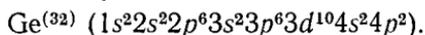
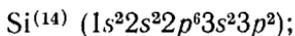
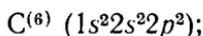
В твердом состоянии при образовании кристалла валентный электрон натрия переходит к атому хлора. В итоге возникают отрицательный ион хлора и положительный ион натрия



У иона хлора в $3p$ -состоянии имеется 6 электронов, что соответствует наиболее устойчивому энергетическому состоянию, т. е. все разрешенные состояния заполнены и на каждом уровне находится по два электрона с противоположно направленными спинами. Из этого уровня 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 $2N$ 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), кремний, германий и серое олово. Электронная структура

этих атомов следующая:



В твердом состоянии у этих элементов в образовании ковалентной связи принимают участие четыре электрона каждого атома. При этом, как следует из рис. 2-4, зоны s и p перекрываются друг с другом, образуя общую зону с числом состояний $8N$. С уменьшением

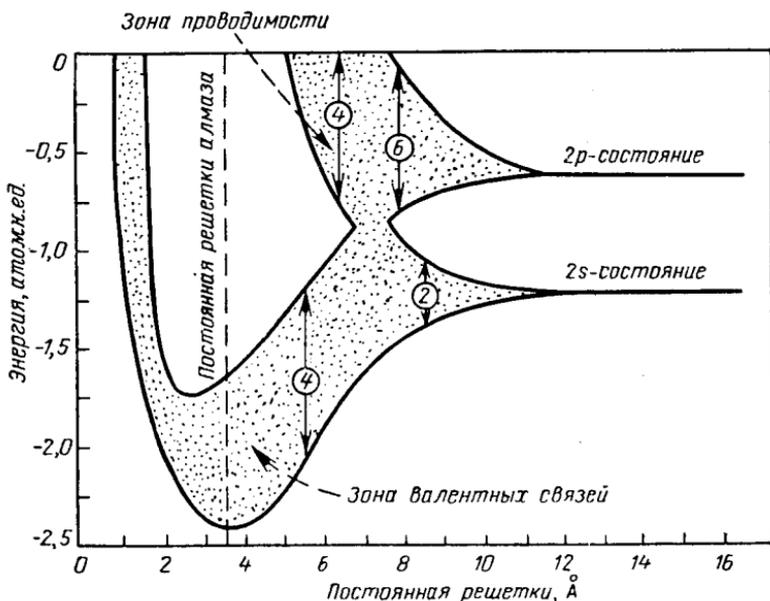


Рис. 2-4. Образование энергетических зон из атомных уровней в алмазе.

межатомного расстояния a эта зона расщепляется на две зоны с $4N$ состояний в каждой. При этом нижняя зона содержит $4N$ заполненных электронами состояний (это валентная зона), а у верхней зоны все $4N$ квантовых состояний свободны — зона проводимости.

2-6. ЗОНЫ БРИЛЛЮЭНА

Как было показано в § 2-4, в разрешенной энергетической зоне кристалла содержится N энергетических состояний, которым соответствуют значения компонент квазиимпульса:

$$p_i = \frac{2\pi\hbar}{L_i} n_j, \quad (2-58)$$

$i = x, y, z; \quad j = 1, 2, 3 \dots$

При этом согласно (2-54) в случае кубической решетки достаточно рассматривать изменение компонент p_i в пределах

$$-\frac{\pi\hbar}{a} \leq p_i < +\frac{\pi\hbar}{a}.$$

Этим значениям квазимпульса в системе координат (p_x, p_y, p_z) соответствует некоторая область, которая называется первой или основной зоной Бриллюэна и для простой кубической решетки имеет форму куба, отсекающего по осям отрезки $\pm \frac{\pi\hbar}{a}$ от начала координат. Для нее объем первой зоны Бриллюэна равен:

$$V_B = \left(\frac{2\pi\hbar}{a}\right)^3 = \frac{h^3}{a^3}. \quad (2-59)$$

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 ($L_x = L_y = L$).^{*} Центр зоны Бриллюэна (начало координат) изображает состояние $n_1 = n_2 = n_3 = 0$,

т. е. $p = 0$. Расстояние между ближайшими точками равно $\frac{2\pi\hbar}{L} = \frac{h}{L}$. Точки, отмеченные кружками, изображают состояния, которые эквивалентны соответствующим состояниям на противоположных гранях зоны Бриллюэна. Например, точки 1 и 2 изображают эквивалентные состояния, имеющие одну и ту же волновую функцию.

Первую зону Бриллюэна можно разбить на кубические ячейки объемом

$$\left(\frac{2\pi\hbar}{L}\right)^3 = \frac{(2\pi\hbar)^3}{V} = \frac{(2\pi\hbar)^3}{a^3 N} = \frac{h^3}{a^3 N}, \quad (2-60)$$

где $V = L^3 = a^3 N$ — объем кристалла, а N — полное число атомов в кристалле. Поскольку объем первой зоны Бриллюэна для кристалла с простой кубической решеткой согласно выражению (2-59) равен $(h/a)^3$, поэтому число элементарных в ней ячеек составляет N , т. е. равно количеству энергетических состояний в зоне. Но в энергетической зоне может располагаться $2N$ электронов, следовательно, и в зоне Бриллюэна может быть $2N$ электронов, а в каждой элементарной ячейке могут находиться два электрона.

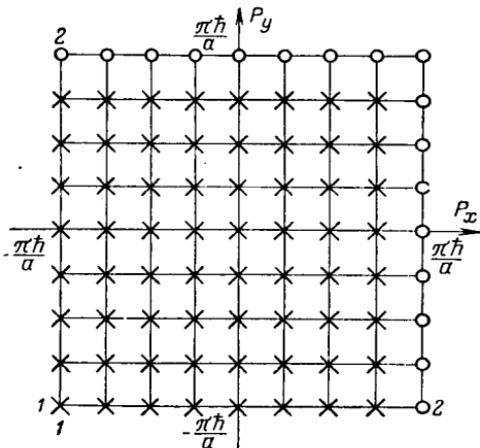
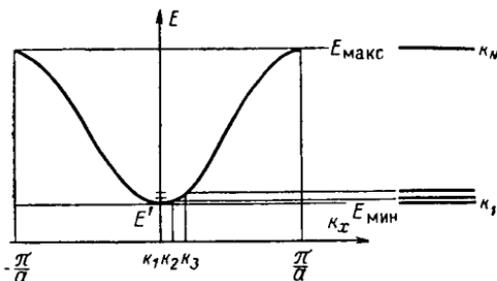


Рис. 2-5. Сечение первой зоны Бриллюэна для простой кубической решетки плоскостью $P_z = 0$.

Вторая и последующая зоны Бриллюэна 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.

2-7. ЗАВИСИМОСТЬ ЭНЕРГИИ ЭЛЕКТРОНОВ ОТ ВОЛНОВОГО ВЕКТОРА У ДНА И ПОТОЛКА ЭНЕРГЕТИЧЕСКОЙ ЗОНЫ

Проиллюстрируем графически зависимость энергии электронов в кристалле от волнового вектора, которая выражается уравнением (2-44). Ее вид зависит



от величины и знака обменного интеграла A . Знак согласно равенству (2-30) величины A определяется перекрытием атомных волновых функций для соответствующего состояния и знаком члена $W(\mathbf{r})$. Член

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

Рис. 2-6. Зависимость $E(k_x)$ для s -состояния.

всегда отрицателен (см. рис. 2-1). Поэтому для s -состояния, имеющего сферическую электронную оболочку, $A < 0$, и выражение (2-43) для компоненты k_x примет вид:

$$E = E' - 2|A| \cos k_x a. \quad (2-61)$$

Полученная зависимость изображена на рис. 2-6 для значений волнового вектора, меняющегося в пределах первой зоны Бриллюэна. Кривая $E(\mathbf{k})$ для зоны s -электронов обращена вогнутостью вверх и для направлений k_y и k_z .

На рис. 2-6 точка $k_x = 0$ определяет состояние с минимальной величиной энергии. Разрешенные уровни энергии соответствуют значениям

$$k_x = \frac{2\pi}{L} n.$$

Точка $k_N = \frac{2\pi}{L} (N - 1)$ соответствует состоянию с максимальным значением энергии.

Для p -состояния, атомные функции Ψ_a которого имеют вид:

$$\frac{x}{r} f(\mathbf{r}), \quad (2-62)$$

перекрывающиеся части функций Ψ_a для атомов с различными координатами, как видно из рис. 2-7, имеют противоположный знак.

Произведение перекрывающихся частей одинакового знака мало, поэтому $A > 0$, и вместо (2-61) получим:

$$E = E'' + 2A \cos \kappa_x a. \quad (2-63)$$

Зависимость энергии от κ_x для p -состояния представлена на рис. 2-8. Здесь при $\kappa_x = 0$ имеется максимум энергии, а κ_N соответствует минимуму энергии.

Для d -состояния получается зависимость, аналогичная s -состоянию, и график $E(\kappa_x)$ имеет вид, подобный представленному на рис. 2-6.

Рассмотрим более детально выражение (2-44) для трехмерного кристалла. Если выбрать для простоты начало отсчета энергии так, чтобы $E_a + C = 0$, получим для s -состояния:

$$E = -2A (\cos \kappa_x a + \cos \kappa_y a + \cos \kappa_z a), \quad (2-64)$$

где $A > 0$.

Обозначим:

$$E_{\min} = -6A = E_d; \quad (2-65)$$

$$E_{\max} = 6A = E_b. \quad (2-65a)$$

Проанализируем равенство (2-64) в окрестности точки $\mathbf{k} = 0$, т. е. около центра первой зоны Бриллюэна.

Разложим функцию $E(\mathbf{k})$ в ряд и ограничимся первыми членами, принимая во внимание, что $\kappa_i a \ll 1$:

$$\begin{aligned} E &= -2A \left\{ \left[1 - \frac{(\kappa_x a)^2}{2} \right] + \left[1 - \frac{(\kappa_y a)^2}{2} \right] + \left[1 - \frac{(\kappa_z a)^2}{2} \right] \right\} = \\ &= -6A + Aa^2 (\kappa_x^2 + \kappa_y^2 + \kappa_z^2) = -6A + Aa^2 \kappa^2. \end{aligned} \quad (2-66)$$

Таким образом, энергия электрона у дна зоны не зависит от направления волнового вектора \mathbf{k} и пропорциональна κ^2 . По аналогии с выражением (2-55) для энергии свободного электрона запишем для электрона в кристалле

$$E = E_d + \frac{\hbar^2}{2m_n^*(d)} \kappa^2, \quad (2-67)$$

где

$$m_n^*(d) = \frac{\hbar^2}{2Aa^2} \quad (2-68)$$

есть эффективная масса электрона у дна энергетической зоны.



Рис. 2-7. Атомные функции p -состояния, описываемые уравнением (2-62).

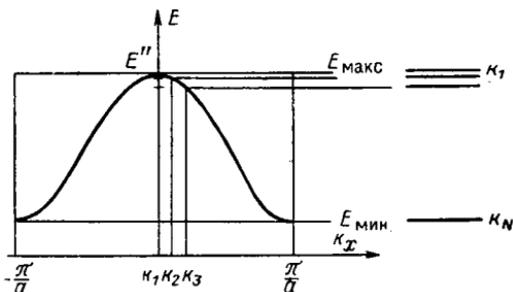


Рис. 2-8. Зависимость $E(\kappa_x)$ для p -состояния.

Рассмотрим формулу (2-64) для состояний, близких к верхнему краю зоны. Введем вектор κ' с компонентами $\kappa'_i = \frac{\pi}{a} - \kappa_i$. При этом у потолка зоны $\kappa_i a \ll 1$ и разложение (2-64) в ряд примет вид:

$$\begin{aligned} E &= -2A [\cos(\pi - \kappa'_x a) + \cos(\pi - \kappa'_y a) + \cos(\pi - \kappa'_z a)] = \\ &= 2A \left\{ \left[1 - \frac{(\kappa'_x a)^2}{2} \right] + \left[1 - \frac{(\kappa'_y a)^2}{2} \right] + \left[1 - \frac{(\kappa'_z a)^2}{2} \right] \right\} = \\ &= 6A - Aa^2 (\kappa')^2 = E_b + \frac{\hbar^2}{2m_n^*(b)} (\kappa')^2. \end{aligned} \quad (2-69)$$

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

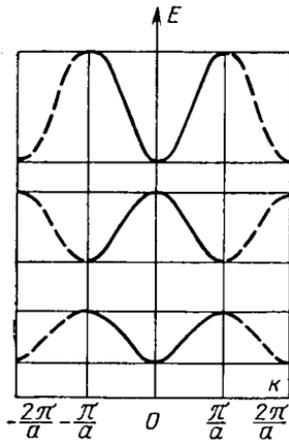


Рис. 2-9. Схематическая зависимость $E(\kappa)$ для одномерной решетки в приближении сильно связанных электронов.

$$m_n^*(b) = -\frac{\hbar^2}{2Aa^2}. \quad (2-70)$$

Поскольку масса не может быть отрицательной, то для состояний у верхнего края зоны вводится понятие дырки, представляющей собой частицу с положительным зарядом, равным по величине заряду электрона, и положительной эффективной массой

$$m_p^*(b) = -m_n^*(b) = \frac{\hbar^2}{2Aa^2}. \quad (2-71)$$

Таким образом, эффективная масса дырки у верхнего края зоны по величине совпадает с эффективной массой электрона у дна этой зоны, она обратно пропорциональна величине обменного интеграла. Но обменный интеграл определяет ширину энергетической зоны. Отсюда следует, что и эффективная масса обратно пропорциональна ширине энергетической зоны. Это означает, что чем шире зона энергии, тем меньше в ней эффективная масса носителей заряда.

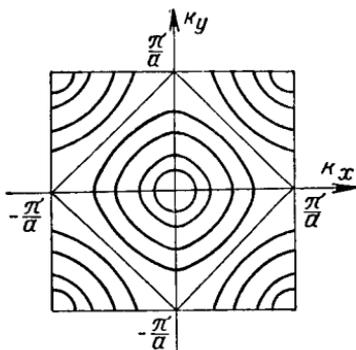
На рис. 2-9 показана зависимость $E(\kappa)$ для трех соседних s , p и d , соответствующая выражению (2-43) для одномерной решетки.

Область изменения κ в пределах $\pm \pi/a$ соответствует первой зоне Бриллюэна. Каждая разрешенная зона содержит N энергетических уровней, среднее расстояние между которыми по порядку величины равно 10^{-22} эв. В точках $\kappa = 0$ и $\kappa = \pm \pi/a$ энергия электрона принимает попеременно минимальное и максимальное значения. Для этих состояний $p = 0$.

Рис. 2-10. Изоэнергетические поверхности для плоской квадратной решетки.

Если положить в выражении (2-64) $E = \text{const}$, мы получим уравнение для изоэнергетических поверхностей в κ -пространстве. Согласно выражениям (2-66) и (2-69) энергия электрона у дна и у потолка зоны не зависит от направления вектора κ , т. е. изоэнергетические поверхности вблизи центра зоны Бриллюэна и около точек $\kappa_i = \pm \pi/a$ имеют сферическую форму.

Вид изоэнергетических поверхностей (кривых) для двумерной решетки приведен на рис. 2-10, из которого следует, что в центре зоны и вблизи углов квадрата изоэнергетические кривые представляют собой окружности, которые искажаются по мере удаления от этих точек.



2-8. ДВИЖЕНИЕ ЭЛЕКТРОНОВ В КРИСТАЛЛЕ ПОД ДЕЙСТВИЕМ ВНЕШНЕГО ЭЛЕКТРИЧЕСКОГО ПОЛЯ

Рассмотрев задачу о движении электрона в периодическом поле кристалла, мы нашли, что энергетический спектр электронов имеет вид квазинепрерывных зон, разделенных промежутками запрещенных значений энергии (рис. 2-3). Поместим кристалл во внешнее постоянное электрическое поле, направленное вдоль оси x , т. е. $\mathcal{E} = \mathcal{E}_x = \text{const}$. Будем считать, что поле мало по сравнению с потенциальной энергией решетки; энергии внешнего поля недостаточно для того, чтобы электрон преодолел запрещенную зону, т. е.

$$Fa \ll \Delta E_g, \quad (2-72)$$

где F — сила, действующая на электрон со стороны внешнего электрического поля; a — параметр решетки; ΔE_g — ширина запрещенной зоны.

Потенциальная энергия электрона во внешнем электрическом поле напряженностью \mathcal{E} будет определяться его координатой x

$$U_1 = -e\mathcal{E}x. \quad (2-73)$$

Уравнение Шредингера для этого случая примет вид:

$$\left\{ -\frac{\hbar^2}{2m} \Delta + [V(\mathbf{r}) + U_1(\mathbf{x})] \right\} \Psi = E' \Psi. \quad (2-74)$$

Для упрощения задачи определения собственных значений энергии E' из (2-74) примем, что потенциал внешнего поля мало изменяется на расстоянии, равном постоянной решетки, так что в области каждого атома можно считать $U_1 = \text{const}$. При этом мы полу-

чаем уравнение Шредингера, подобное уравнению для стационарного случая

$$\left[-\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}) \right] \Psi = (E' - U_1) \Psi. \quad (2-75)$$

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

$$E' = E + U_1, \quad (2-76)$$

где E — энергия электрона в отсутствие поля.

Тогда вместо (2-44) с учетом (2-73) можно написать:

$$E' = E_a + C + 2A (\cos \kappa_x a + \cos \kappa_y a + \cos \kappa_z a) - e\mathcal{E}x. \quad (2-77)$$

Выражение (2-77) свидетельствует о том, что под воздействием внешнего электрического поля энергетические зоны приобретают наклон (рис. 2-11,б).

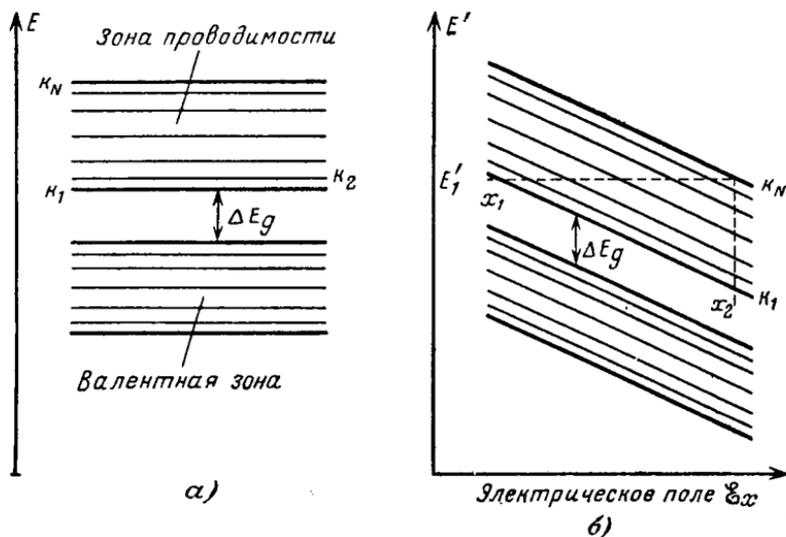


Рис. 2-11. Наклон энергетических зон во внешнем электрическом поле.

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

$$E_{\text{макс}} = E_a + C + 6A - e\mathcal{E}x;$$

$$E_{\text{мин}} = E_a + C - 6A - e\mathcal{E}x.$$

Их разность $E_{\text{макс}} - E_{\text{мин}} = 12A$ не зависит от x и совпадает с шириной зоны для стационарного случая.

В квантовой теории энергия не может зависеть от координат по самой их сути. Это параметр, который может принимать непрерывные или дискретные значения. Поэтому схема, изображен-

ная на рис. 2-11, б, не выражает зависимость энергии от координаты. Она должна восприниматься как характеристика волновой функции.

В отсутствие внешнего поля электрон, находясь на одном из уровней энергетической зоны, обладает определенным значением волнового вектора k , например k_2 на рис. 2-11, а. Это означает, что в стационарном состоянии электрон перемещается по всему кристаллу, имея постоянные значения квазиимпульса и скорости.

При наличии внешнего электрического поля электрон уже не может быть охарактеризован каким-либо определенным значением волнового вектора k . Во внешнем поле полная энергия электрона сохраняется, т. е.

$$E_1' = E + U_1 = \text{const.} \quad (2-78)$$

Но в этом случае электрон при движении по кристаллу переходит с одного уровня зоны на другой (пунктирная кривая на рис. 2-11, б), в результате чего имеет место изменение его волнового вектора, а следовательно, и квазиимпульса. При этом энергия E электрона может меняться от $E_{\text{мин}}$ до $E_{\text{макс}}$. Теперь электрон согласно рис. 2-11, б перемещается не по всему кристаллу, а только в области $\Delta x = x_2 - x_1$, равной

$$\Delta x = \frac{E_{\text{макс}} - E_{\text{мин}}}{e\mathcal{E}} = \frac{12A}{e\mathcal{E}}. \quad (2-79)$$

Таким образом, под действием внешнего электрического поля энергетические зоны идеального кристалла наклоняются, при этом уровни энергии поднимаются, если $U_1 > 0$, и опускаются при $U_1 < 0$. В этом случае электрон движется в некоторой области Δx , ширина которой обратно пропорциональна величине напряженности внешнего поля. При движении в этой области у электрона имеет место изменение волнового вектора и квазиимпульса. Следовательно, их нужно считать теперь зависящими от времени, т. е.

$$k = k(t) \text{ и } p = p(t).$$

Как известно из квантовой механики, движение свободного электрона с волновым вектором 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:

$$\Psi(x, t) = \int_{k - \Delta k}^{k + \Delta k} a_k e^{i(kx - \omega t)} dk, \quad (2-80)$$

где ω — циклическая частота.

Движение волнового пакета характеризуется групповой скоростью $v_{\text{гр}} = \frac{d\omega}{dk}$, равной скорости перемещения, например его

максимума. Координату этого максимума можно найти из условия $\bar{x} - \frac{d\omega}{d\kappa} t = 0$. Отсюда следует, что

$$\frac{\bar{x}}{t} = \frac{d\omega}{d\kappa} = v, \quad (2-81)$$

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

$$v = v_{гп} = \frac{d\omega}{d\kappa}. \quad (2-82)$$

Если воспользоваться соотношением де Бройля, по которому $E = \hbar \omega$, то средняя скорость электрона может быть представлена выражением вида

$$v = \frac{d\omega}{d\kappa} = \frac{1}{\hbar} \frac{dE}{d\kappa}. \quad (2-83)$$

Движение электрона в кристалле описывается волновой функцией Ψ (2-17), которая определяется набором атомных волновых функций с разным значением κ . Поскольку

$$\kappa = \frac{2\pi}{aN} n,$$

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 κ меняется почти непрерывно. В силу этого движение электрона в кристалле можно охарактеризовать волновым пакетом, составленным из блоховских функций. Поэтому выражение (2-83) будет справедливо и для средней скорости движения электрона в кристалле

$$\langle v \rangle = v = \frac{1}{\hbar} \frac{\partial E}{\partial \kappa} = \frac{\partial E}{\partial p}. \quad (2-84)$$

Или для трехмерного случая

$$\mathbf{v} = \frac{1}{\hbar} \text{grad}_{\mathbf{k}} E(\mathbf{k}). \quad (2-85)$$

Вычислим среднюю квантовомеханическую скорость электронов в приближении сильной связи для s -состояния. Из уравнений (2-64) и (2-83) получаем:

$$\mathbf{v} = \frac{2aA}{\hbar} (\mathbf{i} \sin \kappa_x a + \mathbf{j} \sin \kappa_y a + \mathbf{k} \sin \kappa_z a), \quad (2-86)$$

где \mathbf{i} , \mathbf{j} , \mathbf{k} — единичные векторы.

Из (2-86) видно, что средняя скорость электрона зависит как от абсолютной величины, так и от направления волнового вектора κ . Зависимость средней скорости движения электрона $v(\kappa)$ в одномерной решетке представлена на рис. 2-12.

Как видно из рис. 2-12, в точках экстремума энергии ($\kappa = n\pi/a$) значение средней скорости $v = 0$, а в точках внутри первой зоны Бриллюэна находятся максимальные величины v . Средняя скорость всех электронов кристалла в стационарном состоянии

$$\langle v \rangle = \int_{-\pi/a}^{\pi/a} v(\kappa) d\kappa = \frac{2aA}{\hbar} \int_{-\pi/a}^{\pi/a} \sin \kappa a d\kappa = 0. \quad (2-87)$$

Таким образом, в отсутствие внешнего поля средняя скорость электронов равняется нулю. На каждом энергетическом уровне располагается пара электронов (с равными и противоположными по знаку волновыми векторами), движущихся в противоположные стороны.

Рассмотрим случай, когда на электрон действует внешняя сила. Тогда согласно закону сохранения энергии имеем для одномерного движения

$$\frac{dE}{dt} = Fv, \quad (2-88)$$

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

Так как

$$\frac{dE}{dt} = \frac{\partial E}{\partial p} \frac{\partial p}{\partial t}, \quad (2-89)$$

то из сравнения равенств (2-88) и (2-89) с учетом (2-84) следует, что

$$\dot{p} = F. \quad (2-90)$$

Это соотношение по форме совпадает со вторым законом Ньютона, почему p и получил название квазиимпульса.

2-9. ЭФФЕКТИВНАЯ МАССА НОСИТЕЛЕЙ ЗАРЯДА

Если на свободный электрон с массой m действует внешняя сила F , то под ее воздействием электрон приобретает ускорение

$$a = \frac{1}{m} F, \quad (2-91)$$

направленное в сторону действия внешней силы.

Для электрона в кристалле, учитывая равенства (2-84) и (2-90), можно записать:

$$a = \frac{dv}{dt} = \frac{\partial v}{\partial p} \dot{p} = \frac{\partial^2 E}{\partial p^2} F. \quad (2-92)$$

Обобщая равенство (2-92) для трехмерного случая, получаем:

$$\left. \begin{aligned} a_x &= \frac{\partial^2 E}{\partial p_x^2} F_x + \frac{\partial^2 E}{\partial p_x \partial p_y} F_y + \frac{\partial^2 E}{\partial p_x \partial p_z} F_z; \\ a_y &= \frac{\partial^2 E}{\partial p_y \partial p_x} F_x + \frac{\partial^2 E}{\partial p_y^2} F_y + \frac{\partial^2 E}{\partial p_y \partial p_z} F_z; \\ a_z &= \frac{\partial^2 E}{\partial p_z \partial p_x} F_x + \frac{\partial^2 E}{\partial p_z \partial p_y} F_y + \frac{\partial^2 E}{\partial p_z^2} F_z. \end{aligned} \right\} \quad (2-93)$$

В этом случае вектор ускорения не совпадает по направлению с вектором силы.

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 и по аналогии с выражением (2-91) для свободного электрона называется тензором обратной эффективной массы

$$\left\| \frac{1}{m_{ij}^*} \right\| = \begin{vmatrix} \frac{\partial^2 E}{\partial p_x^2} & \frac{\partial^2 E}{\partial p_x \partial p_y} & \frac{\partial^2 E}{\partial p_x \partial p_z} \\ \frac{\partial^2 E}{\partial p_y \partial p_x} & \frac{\partial^2 E}{\partial p_y^2} & \frac{\partial^2 E}{\partial p_y \partial p_z} \\ \frac{\partial^2 E}{\partial p_z \partial p_x} & \frac{\partial^2 E}{\partial p_z \partial p_y} & \frac{\partial^2 E}{\partial p_z^2} \end{vmatrix}. \quad (2-94)$$

Рассмотрим выражение для энергии электрона в малой области окрестности \mathbf{k}_0 . Разложим $E(\mathbf{k})$ в ряд Тейлора, и для области вблизи точки \mathbf{k}_0 ограничимся первыми членами ряда

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \sum_i \left(\frac{\partial E}{\partial k_i} \right)_{\mathbf{k}_0} (k_i - k_{0i}) + \frac{1}{2} \sum_{ij} \left[\left(\frac{\partial^2 E}{\partial k_i \partial k_j} \right)_{\mathbf{k}_0} (k_i - k_{0i}) (k_j - k_{0j}) \right]. \quad (2-95)$$

Здесь i, j независимо пробегает значения x, y, z . Если за \mathbf{k}_0 выбрана точка экстремума энергии, второй член в разложении (2-95) равен нулю. Обозначая

$$\frac{1}{\hbar^2} \left(\frac{\partial^2 E}{\partial k_i \partial k_j} \right)_{\mathbf{k}_0} = \left(\frac{\partial^2 E}{\partial p_i \partial p_j} \right)_{\mathbf{k}_0} = \frac{1}{m_{ij}^*}, \quad (2-96)$$

получаем:

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{\hbar^2}{2} \sum_{ij} \frac{1}{m_{ij}^*} (k_i - k_{0i}) (k_j - k_{0j}). \quad (2-97)$$

Тензор обратной эффективной массы можно сокращенно записать в виде

$$\| m_{ij}^* \|^{\perp} = \begin{bmatrix} m_{xx}^{-1} & m_{xy}^{-1} & m_{xz}^{-1} \\ m_{yx}^{-1} & m_{yy}^{-1} & m_{yz}^{-1} \\ m_{zx}^{-1} & m_{zy}^{-1} & m_{zz}^{-1} \end{bmatrix}. \quad (2-98)$$

Этот тензор симметричен относительно главной диагонали, т. е. $m_{ij}^{-1} = m_{ji}^{-1}$, так как $\partial^2 E / \partial \kappa_i \partial \kappa_j = \partial^2 E / \partial \kappa_j \partial \kappa_i$. Выбрав соответствующую систему координат, можно привести симметричный тензор к диагональному виду, когда отличны от нуля только члены, расположенные на главной диагонали:

$$\|m_{ij}^*\|^{-1} = \begin{bmatrix} m_{xx}^{-1} & 0 & 0 \\ 0 & m_{yy}^{-1} & 0 \\ 0 & 0 & m_{zz}^{-1} \end{bmatrix} = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}. \quad (2-99)$$

При этом выражение (2-97) для энергии электрона вблизи точки κ_0 преобразуется следующим образом:

$$E(\kappa) = E(\kappa_0) + \frac{\hbar^2 (\kappa_x - \kappa_{0x})^2}{2m_1} + \frac{\hbar^2 (\kappa_y - \kappa_{0y})^2}{2m_2} + \frac{\hbar^2 (\kappa_z - \kappa_{0z})^2}{2m_3} \quad (2-100)$$

или

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{(p_x - p_{0x})^2}{2m_1} + \frac{(p_y - p_{0y})^2}{2m_2} + \frac{(p_z - p_{0z})^2}{2m_3}. \quad (2-101)$$

Из равенства (2-101) следует, что поверхности равной энергии в p -пространстве $E(\mathbf{p}) - E(\mathbf{p}_0) = \text{const}$ представляют собой эллипсоиды. Запишем уравнение эллипсоида в канонической форме для постоянного значения $E(\mathbf{p})$:

$$\frac{(p_x - p_{0x})^2}{a^2} + \frac{(p_y - p_{0y})^2}{b^2} + \frac{(p_z - p_{0z})^2}{c^2} = 1, \quad (2-102)$$

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

$$\left. \begin{aligned} a^2 &= 2m_1 [E(\mathbf{p}) - E(\mathbf{p}_0)]; \\ b^2 &= 2m_2 [E(\mathbf{p}) - E(\mathbf{p}_0)]; \\ c^2 &= 2m_3 [E(\mathbf{p}) - E(\mathbf{p}_0)]. \end{aligned} \right\} \quad (2-103)$$

Направление главных осей тензора обратной эффективной массы ориентировано определенным образом относительно осей симметрии кристалла. Если симметрия кристалла такова, что две главные оси тензора физически эквивалентны (например, x и y), соответствующие компоненты эффективной массы равны, т. е. $m_1 = m_2$ и $a = b$. Уравнение изоэнергетической поверхности в этом случае упрощается

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{(p_x - p_{0x})^2 + (p_y - p_{0y})^2}{2m_1} + \frac{(p_z - p_{0z})^2}{2m_3}. \quad (2-104)$$

Изоэнергетическая поверхность, описываемая этим уравнением, является эллипсоидом вращения, у которого ось вращения совпадает с осью P_z . Так как ось вращения соответствует m_3 , то она получила название продольной эффективной массы и обозначается m_l^* . Величина $m_1 = m_2$ называется поперечной эффективной массой и обозначается m_t^* . Если $m_1 < m_3$,

то эллипсоид вращения вытянут вдоль оси вращения. При $m_1 > m_3$ эллипсоид сжат вдоль оси вращения.

Для кристаллов, обладающих кубической симметрией, все три главные оси эквивалентны. При этом $m_1 = m_2 = m_3 = m^*$ и тензор обратной эффективной массы вырождается в скаляр. В этом случае поверхности равной энергии описываются уравнением

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{p^2}{2m^*} \quad (2-105)$$

и представляют собой сферы.

Выражение для эффективной массы носителей заряда в случае кубического кристалла имеет вид:

$$\frac{1}{m^*} = \frac{\partial^2 E}{\partial p^2} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial k^2}, \quad (2-106)$$

т. е. эффективная масса носителей заряда обратно пропорциональна кривизне изоэнергетических поверхностей в k -пространстве.

В этом случае, когда электрон находится в окрестности минимума энергии, т. е. вблизи нижнего края энергетической зоны,

$$\frac{\partial^2 E}{\partial k^2} > 0 \text{ и } m^* = \text{const} > 0. \quad (2-107)$$

При этом согласно уравнению (2-92) получаем:

$$F = m^* a. \quad (2-108)$$

Следовательно, электрон, находящийся у дна энергетической зоны, подчиняется второму закону Ньютона и ведёт себя как частица с положительной массой, равной m^* . Из равенства (2-108) следует, что ускорение электрону сообщает только внешняя сила. Действие поля решетки проявляется в том, что при наличии внешней силы движение электрона определяется не его обычной массой m , а эффективной массой m^* .

В окрестности максимума энергии, т. е. вблизи верхнего края зоны,

$$\frac{\partial^2 E}{\partial k^2} < 0 \text{ и } m^* = \text{const} < 0. \quad (2-109)$$

Ускорение, действующее на электрон, находящийся у верхнего края зоны, направлено противоположно действующей на него внешней силе:

$$\mathbf{a} = - \frac{\mathbf{F}}{m^*}. \quad (2-110)$$

В электрическом поле на электрон действует сила, равная

$$\mathbf{F} = - e\mathcal{E}. \quad (2-111)$$

Для электрона, находящегося вблизи потолка энергетической зоны, у которого $m^* < 0$, ускорение, обусловленное внешней

силой, будет:

$$a = \frac{-|e|\mathcal{E}}{-m^*} = \frac{+e\mathcal{E}}{+m^*}. \quad (2-111a)$$

Такой носитель заряда ведет себя как частица с положительным зарядом и положительной массой и носит название дырки.

По мере удаления от краев зоны зависимость $E(\kappa)$ отклоняется от квадратичного закона. В точках перегиба кривая $E(\kappa)$ при $\kappa = \pm \frac{n\pi}{2a}$ (рис. 2-12) имеет:

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

Из этого следует, что применение соотношения (2-90) для описания движения носителей заряда в кристалле под действием внешнего поля возможно только для носителей, находящихся вблизи дна или потолка энергетической зоны.

На практике наиболее важными оказываются два случая заполнения энергетической зоны.

1. Количество электронов n в зоне много меньше числа возможных квантовых состояний $2N$ (зона почти свободна). Электроны стремятся занять низшие энергетические уровни и располагаются у дна зоны. При этом $m_n^* = \text{const} > 0$.

2. Зона почти заполнена электронами. 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

$$m_p^* = \text{const} > 0.$$

В полупроводниках при температуре абсолютного нуля валентная зона целиком 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 — будет наблюдаться небольшое количество электронов у дна зоны проводимости и дырок у верхнего края валентной зоны. Так как ширина зоны проводимости больше ширины валентной зоны, а величина эффективной массы обратно пропорциональна ширине энергетической зоны, эффективная масса дырки должна быть больше эффективной массы электрона, т. е.

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

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

Существует несколько экспериментальных методов определения эффективной массы носителей заряда. Одним из них является метод циклотронного резонанса.

Рассмотрим его вкратце. Поместим полупроводник в постоянное магнитное поле $\mathbf{B} = \text{const}$. На электрон, движущийся со скоростью \mathbf{v} , будет действовать сила Лоренца

$$\mathbf{F} = e[\mathbf{vB}].$$

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

$$\frac{m^*v^2}{r} = evB, \quad (2-112)$$

откуда следует, что она не зависит от радиуса окружности и выражается формулой

$$\omega_c = \frac{v}{r} = \frac{eB}{m^*}. \quad (2-113)$$

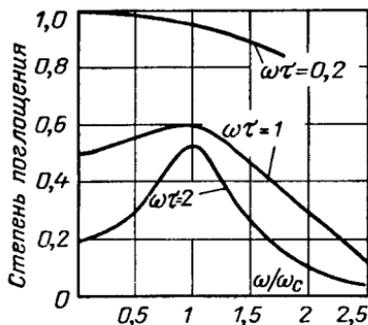


Рис. 2-13. Степень поглощения при циклотронном резонансе в зависимости от напряженности постоянного магнитного поля (в единицах ω/ω_c).

ω_c электрон будет сильно поглощать энергию электрического поля и радиус орбиты его движения будет увеличиваться. Из измерений резонансной частоты поглощения радиоизлучения определяется ω_c , а затем по соотношению (2-113) вычисляется величина эффективной массы электрона (или дырки). Если эта величина анизотропна, то, изменяя ориентацию магнитного и электрического полей относительно кристаллографических осей кристалла, можно изучить форму эллипсоида энергии.

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

$$\tau \gg \frac{1}{\omega}, \quad (2-114)$$

т. е. за время τ электрон должен успеть совершить по меньшей мере один оборот. Чем больше он делает оборотов, тем резче проявляется резонанс. На рис. 2-13 изо-

Если, кроме магнитного поля, приложить слабое высокочастотное электромагнитное поле, колеблющееся в плоскости, перпендикулярной \mathbf{B} , то при приближении частоты поля ω к

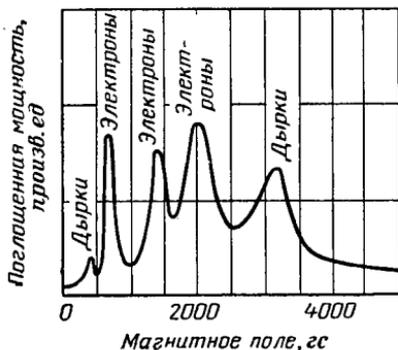


Рис. 2-14. Поглощение при циклотронном резонансе в германии.

бражены кривые поглощения электромагнитного поля в зависимости от ω/ω_c для трех значений $\omega t = 0,2; 1$ 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 $4,2^\circ \text{K}$ — температуре жидкого гелия) и на особо чистых материалах. Обычно в опытах задается частота радиоизлучения, а резонанс достигается изменением магнитного поля. На рис. 2-14 представлены пики поглощения при циклотронном резонансе в германии.

2-10. СТРУКТУРА ЭНЕРГЕТИЧЕСКИХ ЗОН НЕКОТОРЫХ ПОЛУПРОВОДНИКОВ

В качестве примера рассмотрим зонную структуру наиболее изученных в настоящее время полупроводников: германия, кремния и антимонида индия. Германий и кремний имеют кристаллическую структуру типа алмаза, представляющую две гранецентрированные 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. 2-15, где значения составляющих волнового вектора (k_x, k_y, k_z) даны в единицах $2\pi/a$ (a — ребро куба решетки).

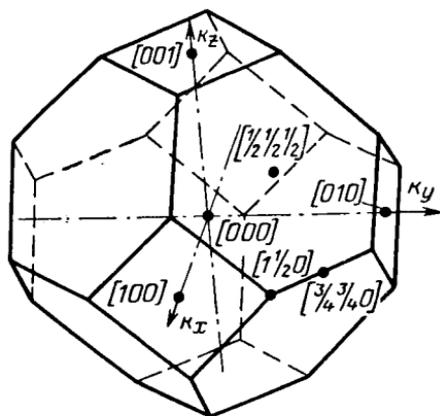


Рис. 2-15. Первая зона Бриллюэна полупроводника типа алмаза.

Теоретический расчет зависимости $E(k)$ встречает большие математические трудности. Для германия и кремния он был проведен лишь для некоторых симметричных точек зоны Бриллюэна, для которых этот расчет сильно упрощается. При этом пришлось решать уравнение 16-й степени. Данные для промежуточных точек были получены интерполированием. Результаты теоретических расчетов уточнялись сравнением с данными экспериментов, в частности, по циклотронному резонансу.

У германия и кремния последняя оболочка не заполнена, в ней в p -состоянии имеется по два электрона с параллельными спинами. Поскольку зона проводимости и валентная зона Ge и Si включают p -состояние, для которого вырождение в кристалле снимается, то в каждой из них имеется по три различные зоны. На рис. 2-16,а они представлены тремя ветвями $E(k)$. Эта зависимость различна для разных направлений.

Одна из ветвей $E(\mathbf{k})$ зоны проводимости как у германия, так и у кремния лежит значительно ниже других. Положение этого абсолютного минимума определяет дно зоны проводимости. Минимумы энергии называют также долинами.

У германия абсолютный минимум лежит в направлении $[111]$, поэтому имеется восемь эквивалентных минимумов энергии, координаты которых в единицах π/a имеют вид $[1/2, 1/2, 1/2]$ и лежат на границе зоны Бриллюэна (на рис. 2-17, а они показаны жирными

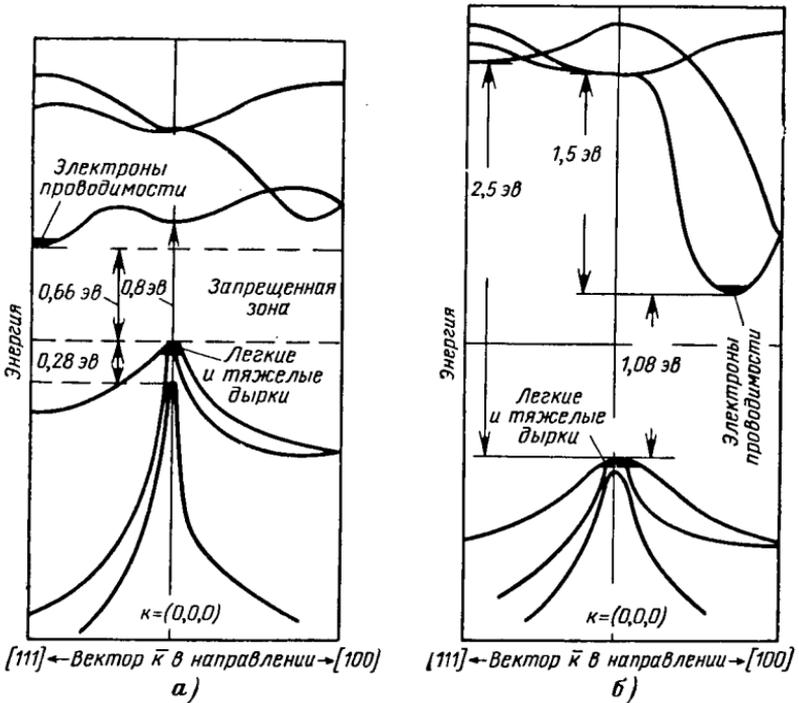


Рис. 2-16. Энергетическая структура германия (а) и кремния (б).

точками). В силу этого на первую зону Бриллюэна приходится половина каждого эллипсоида энергии, следовательно, в германии имеется не восемь, а только четыре полных эллипсоида энергии.

У кремния абсолютный минимум расположен в направлении $[100]$, соответственно имеется шесть эквивалентных минимумов. Изоэнергетические поверхности около абсолютных минимумов (или долин) представляют собой эллипсоиды вращения с осью симметрии, совпадающей с кристаллографическими направлениями $[111]$ в германии (рис. 2-17, а) и с направлениями $[100]$ в кремнии (рис. 2-18, а).

Зависимость энергии от \mathbf{k} вблизи этих минимумов выражается формулой

$$E(\mathbf{k}) = E(\mathbf{k})_0 - \frac{\hbar^2 (\kappa_1 - \kappa_{10})^2}{2m_1} + \frac{\hbar^2 (\kappa_2 - \kappa_{20})^2}{2m_2} + \frac{\hbar^2 (\kappa_3 - \kappa_{30})^2}{2m_3}. \quad (2-115)$$

Значения поперечной m_1 и продольной m_3 эффективных масс из опытов по циклотронному резонансу составляют у германия

$$m_1 = m_2 = m_1^* = 0,082m; \quad m_3 = m_3^* = 1,58m,$$

а у кремния

$$m_1 = m_2 = m_1^* = 0,19m; \quad m_3 = m_3^* = 0,98m.$$

Отношение m_3/m_1 характеризует анизотропию свойств изоэнергетических поверхностей. Для германия это отношение составляет величину 19,3, а у кремния — 5,16.

Как у германия, так и у кремния максимум энергии валентной зоны находится в центре зоны Бриллюэна для всех трех полос (рис. 2-16). В точке $\kappa = 0$ две полосы соприкасаются, т. е. вырождены (одному значению энергии соответствуют различные состояния электрона). Для

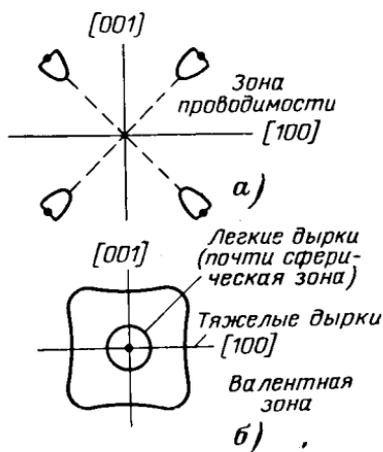


Рис. 2-17. Изоэнергетические поверхности в германии.

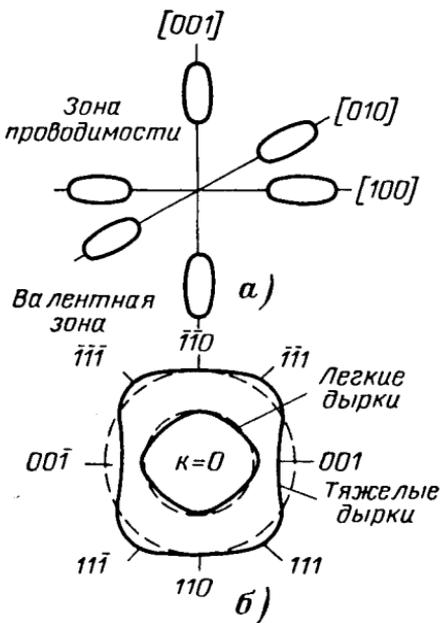


Рис. 2-18. Изоэнергетические поверхности в кремнии.

третьей полосы вырождение частично снимается за счет спин-орбитального расщепления, обусловленного взаимодействием спина электрона с магнитным полем орбитального движения. В результате этого она опускается на величину E_{s0} . Минимальное расстояние между дном зоны проводимости и потолком валентной зоны есть ширина запрещенной зоны. При комнатной температуре у германия ширина запрещенной зоны равна 0,66 эв, а у кремния она составляет 1,08 эв.

Расчет зависимости $E(\kappa)$ для верхних ветвей валентной зоны приводит к выражению

$$E_{1,2}(\kappa) = E(0) - \frac{\hbar^2}{2m} \left[A\kappa^2 \pm \sqrt{B^2\kappa^4 + C^2(\kappa_x^2\kappa_y^2 + \kappa_y^2\kappa_z^2 + \kappa_z^2\kappa_x^2)} \right], \quad (2-116)$$

где A , B и C — безразмерные константы, равные соответственно для германия 13; 8,9 и 10,3 и для кремния 4,1; 1,6 и 3,3. Поверхности равной энергии (2-116) представляют собой гофрированные поверхности (рис. 2-17, б и 2-18, б).

Усреднение зависимости (2-116) по различным направлениям в пространстве позволяет заменить гофрированную поверхность сферической, для которой $E_{1,2}$ может быть представлено выражением

$$E_{1,2} = E(0) - \frac{\hbar^2 \kappa^2}{2m} \left[A \pm \sqrt{B^2 + \frac{C^2}{5}} \right]. \quad (2-117)$$

В этом случае должны существовать тяжелые дырки с эффективной массой

$$m_{pt}^* = \frac{m}{A - \sqrt{B^2 + \frac{C^2}{5}}} \quad (2-118)$$

и легкие дырки

$$m_{pl}^* = \frac{m}{A + \sqrt{B^2 + \frac{C^2}{5}}}. \quad (2-119)$$

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

$$m_{pt}^* = 0,34m; \quad m_{pl}^* = 0,04m;$$

для кремния

$$m_{pt}^* = 0,49m; \quad m_{pl}^* = 0,16m.$$

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

$$E_3(\kappa) = E(0) - E_{s0} - \frac{\hbar^2 \kappa^2}{2m} A, \quad (2-120)$$

где E_{s0} — энергия спин-орбитального взаимодействия. Она равна 0,28 эв для германия и 0,035 эв для кремния.

Поверхности равной энергии третьей валентной зоны согласно (2-120) представляют собой сферы. Поэтому эффективная масса третьего типа дырок является скалярной величиной и равна $m_p^* = m/A$.

Как показывают эксперименты, зона проводимости антимонида индия обладает сферической симметрией и минимум ее находится в центре зоны Бриллюэна. Вблизи минимума энергия подчиняется параболическому закону, а кривизна зоны очень велика. В силу этого значение величины эффективной массы электронов, находящихся на дне этой зоны, очень мало и составляет 0,013 m по данным измерений циклотронного резонанса. Для больших энергий зона проводимости имеет непараболическую форму и кривизна ее уменьшается. Это подтверждают данные рис. 2-19, из которого видно, что энергия не пропорциональна κ^2 .

Валентная зона антимонида индия, зонная структура которого для направления $[111]$ изображена на рис. 2-20, включает три ветви $E(\kappa)$. Зона тяжелых дырок V_1 ($m_{p,r}^* = 0,6 m$) и зона легких дырок V_2 ($m_{p,l}^* = 0,012 m$) вырождены при $\kappa = 0$. Предполагается,

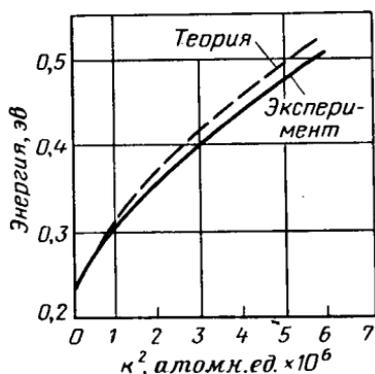


Рис. 2-19. Отклонение формы зоны проводимости в антимониде индия от параболической.

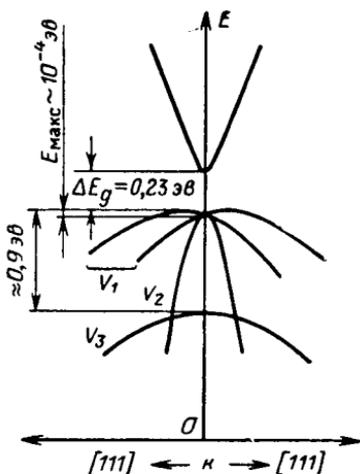


Рис. 2-20. Энергетическая структура антимонида индия.

что максимумы валентной зоны находятся не точно при $\kappa = 0$, а смещены в направлении $[111]$. Третья зона V_3 обусловлена спинорбитальным взаимодействием.

2-11. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ЛОКАЛЬНЫХ УРОВНЕЙ

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:

$$\left[-\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}) + U(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2-121)$$

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

Из предыдущего рассмотрения следует, что квантовые уравнения движения электрона в периодическом поле решетки при наличии внешнего поля аналогичны классическим уравнениям движения свободного электрона, если его массу m заменить в общем случае тензором эффективной массы m^* .

Уравнение Шредингера для свободного электрона имеет вид:

$$-\frac{\hbar^2}{2m} \Delta \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2-122)$$

Запишем уравнение Шредингера для электрона в кристалле соответственно выражению (2-122)

$$-\frac{\hbar^2}{2m^*} \Delta \Psi_0(\mathbf{r}) = E_0 \Psi_0(\mathbf{r}). \quad (2-123)$$

Необходимо выяснить, при каких условиях уравнение (2-123) дает решения, подобные решениям уравнения Шредингера с периодическим потенциалом

$$\left[-\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2-124)$$

Известно, что уравнению (2-124) удовлетворяют блоховские функции $\Psi(\mathbf{r}) = U_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}}$, а энергетический спектр электрона вблизи экстремальной точки ($\mathbf{k} = \mathbf{k}_0$) есть

$$E = E(\mathbf{k}_0) + \frac{\hbar^2}{2} \left[\frac{(\kappa_x - \kappa_{0x})^2}{m_1} + \frac{(\kappa_y - \kappa_{0y})^2}{m_2} + \frac{(\kappa_z - \kappa_{0z})^2}{m_3} \right]. \quad (2-125)$$

В уравнении (2-125) предполагается, что тензор обратной эффективной массы приведен к диагональному виду, и поэтому уравнение (2-123) может быть записано в виде

$$\sum_{i=1}^3 \frac{1}{m_i} \frac{\partial^2 \Psi_0}{\partial x_i^2} + \frac{2E_0}{\hbar^2} \Psi_0 = 0. \quad (2-126)$$

Нетрудно убедиться, что решением этого уравнения будет:

$$\Psi_0 = A e^{i(\alpha x + \beta y + \gamma z)}. \quad (2-127)$$

При этом

$$E_0 = \frac{\hbar^2}{2} \left(\frac{\alpha^2}{m_1} + \frac{\beta^2}{m_2} + \frac{\gamma^2}{m_3} \right). \quad (2-128)$$

Сравнивая выражения (2-125) и (2-128), можно сделать вывод, что если положить $\alpha = (\kappa_x - \kappa_{0x})$, $\beta = (\kappa_y - \kappa_{0y})$, $\gamma = (\kappa_z - \kappa_{0z})$, а начало отсчета энергии вести от экстремальной точки, т. е. $[E(\mathbf{k}_0) = 0]$, то решение уравнения (2-123) дает тот же спектр энергии электрона в кристалле вблизи экстремальной точки ($\mathbf{k} = \mathbf{k}_0$), что и уравнение (2-124).

Таким образом, для решения задачи определения спектра энергии электрона кристалла при наличии внешних полей можно вместо уравнения (2-121) решать уравнение

$$\left[-\frac{\hbar^2}{2m^*} \Delta + U(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}), \quad (2-129)$$

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 справедливы 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 на состояниях с минимальной энергией.

До сих пор мы имели дело с идеальными кристаллами, обладающими периодической структурой. Рассмотрим теперь случай наличия в решетке простейших дефектов 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 парноэлектронной связи, будет взаимодействовать с большим числом атомов кремния, и в результате он оказывается сравнительно слабо связанным с ионом мышьяка силами кулоновского взаимодействия. Его поведение подобно поведению электрона в атоме водорода. Однако при решении задачи о величине потенциала 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 для него необходимо учитывать не обычную массу, а эффективную массу m^* . Во-вторых, взаимодействие между электроном и положительным ионом мышьяка с зарядом $+Ze$ происходит в твердом теле, обладающем диэлектрической проницаемостью ϵ , поэтому в выражение для потенциальной энергии электрона примеси следует ввести величину ϵ . Она будет равна:

$$U(r) = -\frac{Ze^2}{4\pi\epsilon\epsilon_0 r}, \quad (2-130)$$

где ϵ_0 — диэлектрическая постоянная.

Соотношение (2-130) справедливо только для макроскопических точечных зарядов, другими словами, оно справедливо в том случае, если расстояние между взаимодействующими зарядами достаточно велико, т. е. орбита электрона примеси должна охватывать большое число узлов кристаллической решетки.

С учетом сказанного уравнение Шредингера для электрона атома мышьяка, не участвующего в образовании валентных связей, согласно выражению (2-129) можно записать в виде

$$\left[-\frac{\hbar^2}{2m^*} \Delta - \frac{Ze^2}{4\pi\epsilon\epsilon_0 r} \right] \Psi_a = E_n \Psi_a. \quad (2-131)$$

Решив это уравнение по аналогии с уравнением для атома водорода, получим выражение для собственных значений энергии электрона

трона примеси

$$E_n = E_c - \frac{m^* Z^2 e^4}{8h^2 \epsilon^2 \epsilon_0^2 n^2} \quad (n = 1, 2, \dots), \quad (2-132)$$

где энергия электрона отсчитывается от дна зоны проводимости.

Эту формулу можно записать следующим образом:

$$E_n = E_c - \frac{mZ^2 e^4}{8\epsilon_0^2 h^2} \left(\frac{m^*}{m}\right) \frac{1}{\epsilon^2 n^2}. \quad (2-133)$$

Если подставить численное значение e , m , h , ϵ_0 и выразить энергию в электрон-вольтах, то получим:

$$E_n = E_c - \frac{13,52Z^2}{\epsilon^2} \left(\frac{m^*}{m}\right) \frac{1}{n^2} = E_c - \frac{E_d}{n^2} \quad (n = 1, 2 \dots), \quad (2-134)$$

где 13,52 соответствует величине в электрон-вольтах энергии ионизации атома водорода;

$$E_d = - \frac{13,52Z^2}{\epsilon^2} \left(\frac{m^*}{m}\right) \quad (2-135)$$

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

Из (2-134) 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 ($E_c - E_d$) согласно равенству (2-135) в ϵ^2 раз меньше энергии ионизации атома водорода. Кроме того, энергия ионизации зависит от Z^2 . Это означает, что уровень двукратно заряженного иона донорной примеси расположен в запрещенной зоне ниже уровня однократно заряженного иона этой примеси.

Если положить $m^* = 0,25 m$, то в случае германия, у которого $\epsilon = 15,8$, энергия ионизации пятивалентных атомов примеси должна быть порядка 0,01 эв. Как следует из табл. 2-1, величина энергии

Таблица 2-1

Энергия ионизации примесей элементов III и V групп в германии и кремнии¹

Примесь	Донор или акцептор	Энергия ионизации, эв		
		Ge	Si	
B	A	0,0104 (Т)	0,045 (Т)	0,046 (О)
Al	A	0,0102 (Т)	0,057 (Т)	0,067 (О)
Ga	A	0,0108 (Т)	0,065 (Т)	0,071 (О)
In	A	0,0112 (Т)	0,16 (Т)	0,154 (О)
P	Д	0,0120 (Т)	0,044 (Т)	0,046 (О)
As	Д	0,0127 (Т)	0,049 (Т)	0,056 (О)
Sb	Д	0,0096 (Т)	0,039 (Т)	0,043 (О)

¹ Т — величина, определенная термическим методом, О — оптическим.

ионизации, полученная экспериментально, хорошо согласуется с рассчитанным значением.

Расчет радиуса орбиты электрона примеси по водородоподобной модели дает формулу

$$a = \frac{\epsilon_0 \hbar^2}{\pi e^2 m^*} \epsilon n^2 = a_0 \left(\frac{m}{m^*} \right) \epsilon n^2 \quad (n = 1, 2 \dots), \quad (2-136)$$

где $a_0 = \frac{\epsilon_0 \hbar^2}{\pi e^2 m}$ — первая боровская орбита атома водорода.

Радиус первой орбиты электрона донорной примеси в германии согласно выражению (2-136) составляет величину $a_1 = 60 a_0 = 32 \text{ \AA}$. Если учесть, что постоянная решетка германия равна $5,62 \text{ \AA}$, то эта орбита охватывает приблизительно 200 узлов решетки. Это подтверждает возможность рассмотрения движения электрона примеси в среде с постоянной диэлектрической проницаемостью.

Рассмотрим случай акцепторной примеси. Пусть, например, 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 валентной связи с атомом кремния, расположенного рядом с акцепторной примесью.

Таблица 2-2

Энергетические уровни некоторых примесей в германии и кремнии

Примесь	Донор или акцептор	Энергетический уровень, эВ	
		Ge	Si
Li	Д	$E_c - 0,0093$	$E_c - 0,033$
Cu	А	$E_v + 0,04$	$E_v + 0,49$
	А ⁽¹⁾	$E_v + 0,33$	$E_v + 0,24$
Ag	Д	$E_v + 0,13$	
	А	$E_c - 0,29$	
	А ⁽¹⁾	$E_c - 0,04$	
Au	Д	$E_v + 0,05$	$E_v + 0,35$
	А	$E_v + 0,16$	$E_v + 0,54$
Zn	А ⁽¹⁾	$E_c - 0,29$	
	А ⁽²⁾	$E_c - 0,04$	
	Д	$E_v + 0,03$	
Cd	Д ⁽¹⁾	$E_v + 0,09$	
	А		$E_v - 0,31$
	А ⁽¹⁾		$E_v - 0,55$
S	Д	$E_v + 0,05$	
	Д ⁽¹⁾	$E_v + 0,16$	
Se	Д	$E_c - 0,18$	$E_c - 0,18$
Te	Д	$E_c - 0,14$	
Mn	Д	$E_c - 0,10$	
	А	$E_v - 0,16$	
Ni	А ⁽¹⁾	$E_c - 0,37$	
	Д		$E_c - 0,53$
	А	$E_v + 0,22$	
	А ⁽¹⁾	$E_c - 0,30$	

Незавершенная связь — дырка ведет себя как частица с положительным зарядом, что и обеспечивает электронейтральность в этой части кристалла. Решение задачи для этого случая, которая аналогична задаче о движении электрона в кулоновском поле положительного иона, дает собственные значения энергии

$$E_p = E_v + \frac{m^* Z^2 e^4}{8 \epsilon_0^3 \epsilon^2 h^2} \frac{1}{n^2} = E_v + \frac{E_a}{n^2} \quad (n = 1, 2 \dots), \quad (2-137)$$

где
$$E_a = \frac{13,52 Z^2}{\epsilon^2} \left(\frac{m^*}{m} \right), \quad (2-138)$$

т. е. уровень акцепторной примеси находится в запрещенной зоне выше потолка валентной зоны E_v на величину E_a . Согласно расчету она составляет 0,01 эВ для германия (табл. 2-1).

Таким образом, «мелкие» уровни донорной и акцепторной примесей, которые образуют атомы элементов V и III групп в ковалентных полупроводниках типа германия, хорошо описываются в одородной модели. Энергия ионизации этих примесей обратно пропорциональна ϵ^2 и пропорциональна эффективной массе носителей заряда.

Элементы I, II, VI и VIII групп, введенные в полупроводник в качестве легирующих примесей, образуют «глубокие» уровни энергии. В табл. 2-2 даны энергетические уровни некоторых примесей в германии и кремнии.

Эти элементы могут входить в решетку в виде атома (донора Д или акцептора А), 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 образовывать как донорные, так и акцепторные уровни.

Глава третья

СТАТИСТИКА ЭЛЕКТРОНОВ И ДЫРОК В ПОЛУПРОВОДНИКАХ

3-1. ПЛОТНОСТЬ КВАНТОВЫХ СОСТОЯНИЙ

Важнейшая задача статистики 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

$$N(E) = \frac{dZ}{dE}. \quad (3-1)$$

Если вероятность заполнения состояния с энергией E равна $f(E, T)$, то число электронов dn в этих состояниях равно:

$$dn = f(ET) dZ = f(E) N(E) dE. \quad (3-2)$$

Соответственно полное число электронов, для которых возможный интервал энергии лежит в пределах от E_1 до E_2 , будет определяться интегралом

$$n = \int_{E_1}^{E_2} f(E, T) N(E) dE. \quad (3-3)$$

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 уравнению (2-105) энергия электронов у дна зоны может быть записана в виде

$$E = E_c + \frac{p^2}{2m_n^*}, \quad (3-4)$$

где $E_c = E(p_0)$ — энергия электрона у дна зоны проводимости; m_n^* — эффективная масса электрона зоны проводимости.

Выделим шаровой слой, заключенный между двумя изоэнергетическими поверхностями, соответствующими энергиям $E(p) = \text{const}$ и $E(p) + dE = \text{const}$. Объем этого слоя составляет величину

$$dV_p = 4\pi p^2 dp. \quad (3-5)$$

Объем элементарной ячейки зоны Бриллюэна кристалла единичного объема в соответствии с выражением (2-60) равен h^3 . В каждой ячейке могут находиться два электрона с противоположно направленными спинами. С учетом этого число состояний в объеме dV_p равно:

$$dZ = 2 \frac{dV_p}{h^3} = \frac{8\pi}{h^3} p^2 dp. \quad (3-6)$$

Исходя из равенства (3-4), имеем:

$$p^2 = 2m_n^* (E - E_c), \quad (3-7)$$

откуда

$$p = (2m_n^*)^{1/2} (E - E_c)^{1/2}$$

и

$$dp = \frac{1}{2} (2m_n^*)^{1/2} (E - E_c)^{-1/2} dE. \quad (3-8)$$

Подставив равенства (3-6) — (3-8) в соотношение (3-1), получим выражение для плотности квантовых состояний у дна зоны

проводимости, обладающей сферической симметрией:

$$N(E) = 4\pi \left(\frac{2m_n^*}{h^2} \right)^{3/2} (E - E_c)^{1/2}. \quad (3-9)$$

Определим плотность состояний вблизи верхнего края валентной зоны. С этой целью воспользуемся формулами (2-105) и (2-109) для энергии дырок

$$E = E_v - \frac{p^2}{2m_p^*}. \quad (3-10)$$

Здесь E_v — энергия дырок у потолка валентной зоны, а m_p^* — эффективная масса дырки. Из расчетов, аналогичных проделанным выше, будем иметь:

$$N(E) = 4\pi \left(\frac{2m_p^*}{h^2} \right)^{3/2} (E_v - E)^{1/2}. \quad (3-11)$$

Если энергия носителя 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 (3-9) и (3-11) справедливы только для состояний вблизи экстремума энергии, т. е. у дна или потолка зоны.

Подсчитаем теперь плотность квантовых состояний для случая сложной зоны проводимости, когда зависимость $E(p)$ у дна зоны описывается соотношением вида (2-101). Поскольку для дна зоны $p_{0x} = p_{0y} = p_{0z} = 0$, то

$$E = E_c + \frac{p_x^2}{2m_1} + \frac{p_y^2}{2m_2} + \frac{p_z^2}{2m_3}, \quad (3-12)$$

где $1/m_1$, $1/m_2$ и $1/m_3$ — диагональные компоненты тензора обратной эффективной массы. Изоэнергетические поверхности в этом случае представляют собой эллипсоиды, уравнения которых в канонической форме имеют вид:

$$\frac{p_x^2}{a^2} + \frac{p_y^2}{b^2} + \frac{p_z^2}{c^2} = 1, \quad (3-13)$$

где полуоси эллипсоида

$$a_i = [2m_i (E - E_c)]^{1/2}. \quad (3-14)$$

Объем эллипсоида с полуосями a , b , c равен:

$$V = \frac{4\pi}{3} abc = \frac{8\pi}{3} (2m_1 m_2 m_3)^{1/2} (E - E_c)^{3/2}.$$

Объем слоя, заключенного между двумя эллипсоидами равной энергии $E = \text{const}$ и $E + dE = \text{const}$, будет:

$$dV_p = 4\pi (2m_1 m_2 m_3)^{1/2} (E - E_c)^{1/2} dE. \quad (3-15)$$

В объеме dV_p с учетом спина заключено состояний в количестве

$$dZ = 2 \frac{dV_p}{h^3} = 4\pi \left(\frac{2}{h^2} \right)^{3/2} (m_1 m_2 m_3)^{1/2} (E - E_c)^{1/2} dE, \quad (3-16)$$

поэтому выражение для плотности состояний у дна сложной зоны проводимости примет вид:

$$N(E) = 4\pi \left(\frac{2}{\hbar^2}\right)^{3/2} (m_1 m_2 m_3)^{1/2} (E - E_c)^{1/2}. \quad (3-17)$$

Его можно свести к виду, аналогичному выражению (3-9), если положить:

$$m_1 m_2 m_3 = m_{dn}^{*3}, \quad (3-18)$$

где m_{dn}^* называется эффективной массой для плотности состояний. При этом, как и для простой зоны, получим:

$$N(E) = 4\pi \left(\frac{2m_{dn}^*}{\hbar^2}\right)^{3/2} (E - E_c)^{1/2}. \quad (3-19)$$

Таким образом, эллипсоидальную поверхность равной энергии зоны проводимости с точки зрения плотности квантовых состояний можно заменить сферой, которой соответствует эффективная масса плотности состояний, вычисляемая согласно формуле (3-18).

Найдем плотность состояний для кремния и германия.

Кремний. Как показано на рис. 2-16, минимум энергии в зоне проводимости кремния расположен в направлении [100] на расстоянии от точки $\kappa = 0$, равном приблизительно 75% величины $2\pi/a$. Всего в первой зоне Бриллюэна располагается шесть минимумов, и для энергии, несколько превышающей величину E_c , образуется шесть эллипсоидов равной энергии, вытянутых вдоль главных кристаллографических осей (рис. 2-18, а). В соответствии с этим плотность состояний, которая выражается для одного эллипсоида равенством (3-17), увеличится в 6 раз. И если учесть, что для кремния $m_1 = m_2$, то

$$N(E) = 6 \cdot 4\pi \left(\frac{2}{\hbar^2}\right)^{3/2} (m_1^2 m_3)^{1/2} (E - E_c)^{1/2}, \quad (3-20)$$

а эффективная масса плотности состояния с учетом значения $m_1 = 0,19 m$, $m_3 = 0,98 m$ будет:

$$m_{dn}^* = 6^{2/3} (m_1^2 m_3)^{1/3} = 1,08 m. \quad (3-21)$$

Следовательно, у кремния все шесть эллипсоидов изоэнергетической поверхности зоны проводимости можно заменить одной сферической поверхностью с эффективной массой для плотности состояния, равной $1,08 m$.

В точке $\kappa = 0$ валентная зона кремния дважды вырождена, и поверхности равной энергии имеют сложный вид, описываемый формулой (2-116). Однако их можно аппроксимировать двумя сферическими поверхностями вида (2-117), которым соответствуют эффективные массы тяжелых m_{pT}^* и легких дырок m_{pL}^* . В этом случае плотность состояний будет определяться суммой плотности состо-

яний в зонах тяжелых и легких дырок:

$$N(E) = 4\pi \left(\frac{2}{h^2} \right)^{3/2} (m_{\text{рт}}^{*3/2} + m_{\text{рл}}^{*3/2}) (E_v - E)^{1/2}. \quad (3-22)$$

Изоэнергетические поверхности обеих зон можно заменить одной приведенной сферой с плотностью состояний

$$N(E) = 4\pi \left(\frac{2m_{\text{др}}^*}{h^2} \right)^{3/2} (E_v - E)^{1/2}, \quad (3-23)$$

для которой эффективная масса для плотности состояний равна:

$$m_{\text{др}}^* = (m_{\text{рт}}^{*3/2} + m_{\text{рл}}^{*3/2})^{2/3} = 0,59m. \quad (3-24)$$

Германий. В зоне проводимости германия имеется 8 минимумов энергии, расположенных на границе зоны Бриллюэна в точках $[1/2, 1/2, 1/2]$ (рис. 2-15). Соответствующие поверхности равной энергии изображены на рис. 2-17, а, где показаны четыре из восьми эллипсоидов, рассеченные пополам гранями зоны Бриллюэна. Две противоположные половины эллипсоидов (рис. 2-17, а) могут рассматриваться как одна целая эллипсоидальная поверхность. Таким образом, в основной зоне Бриллюэна германия расположены четыре полных эллипсоида, для которых $m_1 = m_2$. Следовательно, плотность состояний у потолка зоны проводимости германия будет определяться выражением (3-19) с эффективной массой для плотности состояний, равной

$$m_{\text{дн}}^* = (4^2 m_1^2 m_2)^{1/3} = 0,56m. \quad (3-25)$$

Структура валентной зоны германия подобна структуре зоны кремния (рис. 2-16). Эффективная масса для плотности состояния дырок, характеризующая плотность состояния у потолка валентной зоны германия, будет:

$$m_{\text{др}}^* = (m_{\text{рт}}^{*3/2} + m_{\text{рл}}^{*3/2})^{2/3} = 0,37m. \quad (3-26)$$

3-2. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ ФЕРМИ — ДИРАКА

Для определения числа частиц, имеющих энергию в заданном интервале, помимо плотности 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 полуцелым спином, подчиняющихся принципу Паули, справедливо распределение Ферми — Дирака

$$f_0(E) = \frac{1}{\frac{E - F}{k_0 T} + 1}, \quad (3-27)$$

где k_0 — постоянная Больцмана; T — абсолютная температура; F — энергия Ферми или химический потенциал, т. е. работа, которую необходимо затратить для изменения числа частиц в системе на одну единицу.

Рассмотрим вид функции распределения Ферми — Дирака при различных температурах (рис. 3-1).

Из соотношения (3-27) 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 выражения (3-27) для значения энергии, равной Ферми ($E = F$), имеем $f_0 = 0,5$. Таким образом, уровень Ферми есть энергетический уровень, вероятность заполнения которого при температуре, отличной от абсолютного нуля, равна 0,5. При $T > 0$ часть электронов за счет теплового движения смогут перейти в состояния с $E > F$, и соответственно часть состояний, лежащих ниже уровня Ферми, оказываются свободными. Число частиц, перешедших на более высокие уровни, равно количеству образовавшихся свободных состояний в области $E < F$. Это означает, что площади, заштрихованные на рис. 3-1, равны друг другу.

Произведем оценку области изменения функции распределения $f_0(E)$ для случая $T > 0$, изображенного на рис. 3-1. Для этого подсчитаем значение $f_0(E)$ для разных величин энергии. Для энергий, отличающихся от F на $\pm k_0 T$, величина $f_0(E)$ составляет $(1 + e)^{-1} = 0,27$ и $(1 + \frac{1}{e})^{-1} = 0,73$. При изменении энергии на $\pm 2k_0 T$ величина f_0 равна 0,118 и 0,882, а при изменении на $\pm 3k_0 T$ соответствует 0,048 и 0,954. Из этих данных следует, что вероятность заполнения состояний заметно отличается от единицы или нуля лишь в пределах $(2 \div 3) k_0 T$ вблизи значения $E = F$ (рис. 3-1).

Функция распределения Ферми — Дирака характеризует вероятность заполнения данного состояния электроном. Вероятность того, что при тепловом равновесии в состоянии с энергией E электрон отсутствует, т. е. оно занято дыркой, будет равна:

$$f_{op}(E) = 1 - f_0(E) = 1 - \frac{1}{e^{\frac{E-F}{k_0 T}} + 1} = \frac{1}{e^{\frac{F-E}{k_0 T}} + 1} \quad (3-28)$$

Следовательно, функция распределения для дырок аналогична функции распределения для электронов, если отсчитывать энергию дырок от уровня Ферми в противоположную сторону по срав-

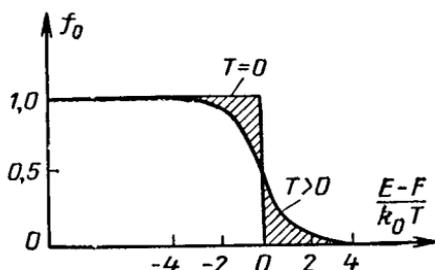


Рис. 3-1. Вид функции распределения Ферми — Дирака при $T = 0$ и $T > 0$.

нению с направлением отсчета энергий для электронов. Полупроводники, у которых свободные носители заряда описываются функцией Ферми — Дирака, называются вырожденными.

Для электронов, находящихся в состояниях с энергией $E - F \gg k_0 T$, выражение для f_0 принимает вид:

$$f_0(E) = e^{\frac{F-E}{k_0 T}} = C e^{-\frac{E}{k_0 T}}, \quad (3-29)$$

т. е. совпадает с функцией распределения Максвелла — Больцмана для частиц, подчиняющихся классическим законам. Если носители заряда в полупроводнике подчиняются статистике Больцмана, то полупроводник называется невырожденным.

3.3. СТЕПЕНЬ ЗАПОЛНЕНИЯ ПРИМЕСНЫХ УРОВНЕЙ

Рассмотрим полупроводник, содержащий одновалентную донорную примесь с концентрацией N_d . Это соответствует случаю, например, когда один из узлов кристаллической решетки кремния занят

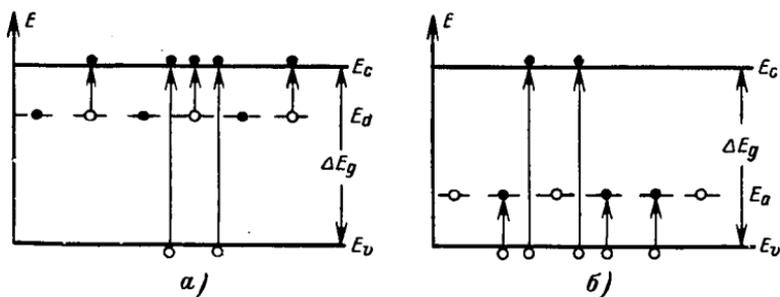


Рис. 3-2. Электронный (а) и акцепторный (б) полупроводники.

атомом мышьяка. При этом пятый валентный электрон атома может быть легко ионизирован. Примесным электронам соответствует энергетический уровень, расположенный ниже дна зоны проводимости на величину E_d (рис. 3-2, а).

Поскольку у донорной примеси имеется только один электрон, который может принимать участие в проводимости, то полное число энергетических состояний донорной примеси должно быть равно количеству атомов введенной примеси на единицу объема кристалла, т. е. равно N_d .

Пусть число электронов, находящихся на уровне донорной примеси, будет n_d . Тогда концентрация ионизированных донорных атомов p_d , образовавшихся в результате переходов электронов с донорных уровней в зону проводимости за счет тепловой энергии и имеющих положительный заряд, будет равно:

$$p_d = N_d - n_d. \quad (3-30)$$

Если бы на примесном донорном уровне согласно принципу Паули могли расположиться два электрона с противоположно на-

правленными спинами, то вероятность его заполнения определялась бы функцией Ферми—Дирака (3-27), в которой вместо E следовало поставить энергию электрона на уровне примеси. Однако на уровне E_d находится только один электрон, поэтому нейтральное состояние примеси имеет вдвое больший статистический вес, характеризующийся числом совпадающих уровней энергии, по сравнению с ионизированным состоянием. Тогда, исходя из принципа Больцмана, можно ожидать, что

$$p_d : n_d = 1 : 2e^{-\frac{E_d - F}{k_0 T}}. \quad (3-31)$$

Используя (3-30), это равенство можно записать в виде

$$\frac{N_d - n_d}{n_d} = \left(2e^{-\frac{E_d - F}{k_0 T}} \right)^{-1}, \quad (3-32)$$

откуда найдем концентрацию электронов, находящихся на донорном уровне, которая будет равна:

$$n_d = \frac{N_d}{\frac{1}{2} e^{\frac{E_d - F}{k_0 T}} + 1}, \quad (3-33)$$

а концентрация положительных ионов донорной примеси на основании соотношений (3-31) и (3-33) будет выражаться уравнением вида

$$p_d = \frac{N_d}{\frac{F - E_d}{2e^{\frac{k_0 T}{k_0 T}} + 1}}. \quad (3-34)$$

Предэкспоненциальный множитель в выражении (3-33) в общем случае можно записать через β^{-1} . Тогда вероятность нахождения электрона на донорном уровне с энергией E_d будет определяться выражением

$$f_0(E_d) = \frac{1}{\frac{1}{\beta} e^{\frac{E_d - F}{k_0 T}} + 1}, \quad (3-35)$$

а функция распределения для положительных ионов донорной примеси на основании равенств (3-28) и (3-35) будет:

$$f_{p_0}(E_d) = \frac{1}{\frac{F - E_d}{2e^{\frac{k_0 T}{k_0 T}} + 1}} = \frac{1}{\beta e^{\frac{F - E_d}{k_0 T}} + 1}. \quad (3-36)$$

Из выражения (3-35) следует, что для одновалентной донорной примеси, для которой примесный уровень двукратно вырожден, фактором (степенью) спинового вырождения является $\beta = 2$.

Рассмотрим теперь акцепторный полупроводник, например кремний, легированный бором. Пусть концентрация введенной примеси равна N_a . Энергетическая схема такого полупроводника представлена на рис. 3-2, б.

Нейтральный атом бора с соседними атомами кремния образует три парноэлектронные связи, четвертая связь одного из четырех соседних атомов кремния будет не завершена; располагаясь около атома бора, она ведет себя как положительная дырка. В эту незавершенную связь может перейти электрон от основного вещества. Для этого потребуется энергия, равная E_a . В результате образуется свободная дырка, а атом бора превращается в отрицательно заряженный ион бора. Так как атом акцепторной примеси для укомплектования парной связи захватывает электрон из валентной зоны, при этом дополнительный электрон может быть захвачен двояким образом с учетом направления спина, то степень вырождения акцепторного уровня $\beta = 2$. Поэтому концентрация дырок p_a на акцепторной примеси при данной температуре T будет определяться соотношением вида

$$p_a = \frac{N_a}{\frac{1}{2} e^{\frac{F - E_a}{k_0 T}} + 1} = \frac{N_a}{\beta^{-1} e^{\frac{E_a - F}{k_0 T}} + 1}, \quad (3-37)$$

а концентрация электронов на акцепторной примеси n_a (или концентрация отрицательных ионов) соответственно будет равна:

$$n_a = N_a - p_a = \frac{N_a}{e^{\frac{E_a - F}{2e k_0 T}} + 1} \leftarrow \frac{N_a}{\beta e^{\frac{E_a - F}{k_0 T}} + 1}. \quad (3-38)$$

3-4. КОНЦЕНТРАЦИЯ ЭЛЕКТРОНОВ И ДЫРОК

Рассмотрим процесс образования свободных носителей заряда под воздействием тепла в примесном полупроводнике, например в донорном, энергетическая схема которого изображена на рис. 3-2, а.

При температуре абсолютного нуля электроны занимают состояния с минимальным значением энергии. Другими словами, при $T = 0$ все состояния донорной примеси и в валентной зоне заняты электронами, поэтому в зоне проводимости нет электронов. При повышении температуры колебания атомов кристалла становятся более интенсивными. Это может привести к отрыву электрона от донорной примеси и даже к разрыву валентных связей основного вещества, и, следовательно, появятся свободные электроны и дырки. *Свободные носители заряда, возникшие за счет энергии тепловых колебаний решетки и находящиеся с ней в термодинамическом равновесии, называются равновесными, тепловыми или темновыми, а процесс их образования носит названия тепловой генерации.*

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. 3-2, а показаны стрелками. *Обратный процесс перехода электрона из зоны проводимости на уровень донорной примеси или в валентную зону называется рекомбинацией.*

В случае акцепторного полупроводника равновесные электроны и дырки появляются в результате теплового перехода электрона из валентной зоны на уровень акцепторной примеси и в зону проводимости (рис. 3-2, б).

Определим при термодинамическом равновесии для примесного полупроводника значения концентрации электронов в зоне проводимости 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 n_0 можно найти, если в равенство (3-3) подставить плотность квантовых состояний для случая изоэнергетических поверхностей (3-9) и воспользоваться функцией распределения Ферми—Дирака (3-27). 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 бесконечности. В результате будем иметь:

$$n_0 = \int_{E_c}^{\infty} f_0(E) N(E) dE = \int_{E_c}^{\infty} \frac{4\pi \left(\frac{2m_n^*}{h^2}\right)^{3/2} (E - E_c)^{1/2} dE}{e^{\frac{E-F}{k_0T}} + 1}. \quad (3-39)$$

Введем безразмерные величины

$$\frac{E - E_c}{k_0T} = \varepsilon; \quad \frac{F - E_c}{k_0T} = \eta. \quad (3-40)$$

Величину η называют приведенным уровнем Ферми. При этом равенство (3-39) преобразуется к виду

$$n_0 = 4\pi \left(\frac{2m_n^*}{h^2}\right)^{3/2} (k_0T)^{3/2} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} = N_c F_{1/2}(\eta). \quad (3-41)$$

Здесь

$$N_c = 2 \left(\frac{2\pi m_n^* k_0T}{h^3}\right)^{3/2} \quad (3-42)$$

и называется эффективной плотностью состояний в зоне проводимости;

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} \quad (3-43)$$

— интеграл Ферми порядка $1/2$.

Эффективная плотность состояний N_c зависит от температуры. Если в (3-42) подставить численные значения универсальных констант, то будем иметь:

$$N_c = 4,82 \cdot 10^{15} \left(\frac{m_n^*}{m}\right)^{3/2} T^{3/2} = 2,5 \cdot 10^{19} \left(\frac{m_n^*}{m}\right)^{3/2} \left(\frac{T}{300}\right)^{3/2}.$$

Найдем количество дырок в валентной зоне. В соответствии с выражениями (3-11) и (3-28) получаем:

$$p_0 = \int_{-\infty}^{E_v} f_{0p}(E) N(E) dE = 4\pi \left(\frac{2m_p^*}{h^2} \right)^{3/2} \int_{-\infty}^{E_v} \frac{(E_v - E)^{1/2} dE}{e^{\frac{F-E}{k_0T}} + 1}. \quad (3-44)$$

В этом равенстве нижний предел интегрирования $E_{v \text{ мин}}$ заменили на $-\infty$, учитывая резкую зависимость f_{0p} от энергии. Обозначим

$$\frac{E_v - E}{k_0T} = \varepsilon_p; \quad \frac{E_c - E_v}{k_0T} = \varepsilon_i; \quad \frac{F - E}{k_0T} = \eta + \varepsilon_i + \varepsilon_p, \quad (3-45)$$

где $E_c - E_v = \Delta E_g$ — ширина запрещенной зоны.

Теперь выражение для концентрации дырок примет вид:

$$p_0 = \frac{4}{\sqrt{\pi}} \left(\frac{2\pi m_p^* k_0T}{h^2} \right)^{3/2} \int_0^{\infty} \frac{\varepsilon_p^{1/2} d\varepsilon_p}{e^{\varepsilon_p + \eta + \varepsilon_i} + 1} = N_v F_{1/2}(-\eta - \varepsilon_i), \quad (3-46)$$

где

$$N_v = 2 \left(\frac{2\pi m_p^* k_0T}{h^2} \right)^{3/2}, \quad (3-47)$$

есть эффективная плотность состояний в валентной зоне;

$$F_{1/2}(-\eta - \varepsilon_i) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon_p^{1/2} d\varepsilon_p}{e^{\varepsilon_p + \eta + \varepsilon_i} + 1}, \quad (3-48)$$

есть интеграл Ферми для валентной зоны.

Если положить $m_n^* = m_p^* = m$, то при $T = 300^\circ \text{C}$ $N_c = N_v = = 2,5 \cdot 10^{19} \text{ см}^{-3}$. Это составляет приблизительно 1/2000 числа атомов в 1 см^3 .

Если ввести обозначения $E_c - E_d/k_0T = \varepsilon_d$, то на основании выражения (3-35) концентрация электронов на уровне донорной примеси будет равна:

$$n_d = f_0(E_d) N_d = \frac{N_d}{\beta^{-1} e^{-\varepsilon_d - \eta} + 1}, \quad (3-49)$$

а количество положительных ионов донорной примеси согласно выражению (3-36) составит величину

$$P_d = f_{0p}(E_d) N_d = \frac{N_d}{\beta e^{\varepsilon_d + \eta} + 1}. \quad (3-50)$$

Найдем концентрацию дырок p_a и концентрацию электронов n_d на акцепторной примеси. Если обозначить $(E_c - E_a)/k_0T = = \varepsilon_a$, то они на основании соотношений (3-37) и (3-38) будут выра-

жаться соотношениями вида

$$p_a = \frac{N_a}{\beta^{-1} e^{\eta + \epsilon} a + 1}; \quad (3-51)$$

$$n_a = \frac{N_a}{\beta e^{-\eta - \epsilon} a + 1}. \quad (3-52)$$

3-5. ПРИМЕСНЫЙ ПОЛУПРОВОДНИК

Для того чтобы подсчитать равновесные концентрации электронов n_0 и дырок p_0 , необходимо вычислить значения интегралов Ферми (3-43) и (3-48), что во многих конкретных случаях представляет значительные трудности. Но для интеграла Ферми существуют приближенные аналитические выражения и составлены таблицы.

В зависимости от η интеграл Ферми (3-43)

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{1/2} d\epsilon}{e^{\epsilon - \eta} + 1}$$

можно выразить следующим образом.

1. При $-\infty < \eta < -1$

$$F_{1/2}(\eta) = e^{\eta}. \quad (3-53)$$

Такое приближение соответствует статистике Больцмана, а полупроводник является невырожденным. Это классическое приближение, для которого

$$n_0 = N_c e^{\eta}, \quad (3-54)$$

на рис. 3-3 изображено кривой 1.

2. При $5 < \eta < \infty$

$$F_{1/2}(\eta) = \frac{4}{3\sqrt{\pi}} \eta^{3/2}. \quad (3-55)$$

Это приближение справедливо для случая полного вырождения и на рис. 3-3 характеризуется кривой 2.

3. При $-1 < \eta < 5$ интеграл Ферми можно приближенно записать в виде

$$F_{1/2}(\eta) = \frac{1}{0,25 + e^{-\eta}}. \quad (3-56)$$

При этих условиях свойства полупроводника являются переходными от невырожденного к полностью вырожденному. Расчет ин-

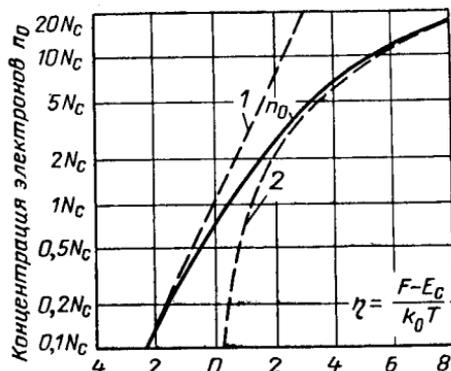


Рис. 3-3. Зависимость концентрации электронов проводимости от положения приведенного уровня Ферми.

теграла Ферми по формуле (3-56) дает значительную погрешность, а поэтому целесообразно пользоваться специальными таблицами для величины $F_{1/2}(\eta)$.

На рис. 3-3 кривая n_0 представляет собой истинное значение концентрации равновесных электронов проводимости, подсчитанное по соотношению (3-41).

Рассмотрим более подробно оба крайних случая.

Невырожденный примесный полупроводник. Принято считать, что полупроводник не вырожден, если

$$\eta = \frac{F - E_c}{k_0 T} < -1 \quad (\text{см. рис. 3-3}) \quad \text{или} \quad F < E_c - k_0 T. \quad (3-57)$$

Это означает, что в невырожденном донорном полупроводнике уровень Ферми лежит ниже дна зоны проводимости по крайней мере на $k_0 T$. 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.

$$f_0 = \frac{1}{e^{\frac{E-F}{k_0 T}} + 1} \approx e^{-\frac{E-F}{k_0 T}}. \quad (3-58)$$

Из равенства (3-58) следует, что в невырожденном полупроводнике носители заряда подчиняются статистике Больцмана. В соответствии с этим интеграл Ферми запишется в виде

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} \approx \frac{2}{\sqrt{\pi}} e^{\eta} \int_0^{\infty} \varepsilon^{1/2} e^{-\varepsilon} d\varepsilon = e^{\eta}. \quad (3-59)$$

Равновесные концентрации электронов n_0 и дырок p_0 для случая невырожденного полупроводника соответственно равны:

$$n_0 = N_c F_{1/2}(\eta) = N_c e^{\eta} = N_c e^{-\frac{E_c - F}{k_0 T}}; \quad (3-60)$$

$$p_0 = N_v F_{1/2}(-\eta - \varepsilon_i) = N_v e^{-\eta - \varepsilon_i} = N_v e^{\frac{E_v - F}{k_0 T}}. \quad (3-61)$$

В собственном полупроводнике концентрации свободных электронов и дырок равны между собой. Если для собственного полупроводника обозначить приведенный уровень Ферми через η_i , а $\frac{E_i - E_v}{k_0 T} = \varepsilon_i$, то

$$n_0 = p_0 = n_i = N_c e^{\eta_i} = N_v e^{-\eta_i - \varepsilon_i}. \quad (3-62)$$

Выразим теперь равновесные концентрации электронов n_0 (3-60) и дырок p_0 (3-61) для невырожденного полупроводника через n_i :

$$n_0 = n_i e^{\eta - \eta_i}; \quad p_0 = n_i e^{-\eta + \eta_i}. \quad (3-63)$$

Отсюда следует:

$$n_0 p_0 = n_i^2, \quad (3-64)$$

т. е. в невырожденном примесном полупроводнике произведение концентраций свободных электронов и дырок при тепловом равновесии есть постоянная величина, равная квадрату концентрации носителей в собственном полупроводнике при данной температуре.

Для дырок по аналогии с электронами интеграл Ферми можно записать:

$$F_{1/2}(-\eta - \varepsilon_i) = \begin{cases} e^{-\eta - \varepsilon_i} & \text{при } -\infty < -(\eta + \varepsilon_i) < -1, \\ \frac{1}{0,25 + e^{-\eta - \varepsilon_i}} & \text{при } -1 < -(\eta + \varepsilon_i) < 5, \\ \frac{4}{3\sqrt{\pi}} (-\eta - \varepsilon_i)^{3/2} & \text{при } 5 < -(\eta + \varepsilon_i) < \infty. \end{cases} \quad (3-65)$$

При $-\eta - \varepsilon_i < -1$ или $F > E_v + k_0 T$ дырочный полупроводник не вырожден и уровень Ферми лежит выше потолка валентной зоны не менее чем на $k_0 T$, а равновесные концентрации электронов и дырок определяются равенствами (3-60) и (3-61). Если $-\eta - \varepsilon_i > 5$ то $F < E_v - 5k_0 T$. Следовательно, акцепторный полупроводник является полностью вырожденным, если уровень Ферми лежит в валентной зоне на $5 k_0 T$ ниже E_v .

Полностью вырожденный примесный полупроводник. Донорный полупроводник считается полностью вырожденным при $\eta = \frac{F - E_c}{k_0 T} > 5$ (рис. 3-3), т. е. когда уровень Ферми расположен в зоне проводимости по крайней мере на $5 k_0 T$ выше E_c . При температуре, близкой к абсолютному нулю ($F - E_c \gg k_0 T$), размытием функции $f_0(E)$ можно пренебречь (см. рис. 3-1), заменив ее ступенчатой зависимостью $f_0 = 1$ при $E < F_0$ и $f_0 = 0$ при $E > F_0$, где F_0 — уровень Ферми при $T = 0$.

В этом случае в выражении (3-43) знаменатель равен единице и равновесная концентрация электронов в вырожденном донорном полупроводнике будет равна:

$$\begin{aligned} n_0 &= N_c \frac{2}{\sqrt{\pi}} \int_0^{\eta_0} e^{\varepsilon/2} d\varepsilon = \frac{4}{3\sqrt{\pi}} N_c \eta_0^{3/2} = \frac{4}{3\sqrt{\pi}} N_c \left(\frac{F_0 - E_c}{k_0 T} \right)^{3/2} = \\ &= \frac{8\pi}{3} \left(\frac{2m_n^*}{h^2} \right)^{3/2} (F_0 - E_c)^{3/2}, \end{aligned} \quad (3-66)$$

а концентрация равновесных дырок p_0 определяется соотношением (3-61).

Равновесная концентрация дырок в вырожденном акцепторном полупроводнике в соответствии с выражениями (3-46) и (3-65) запишется:

$$p_0 = N_c \frac{4}{3\sqrt{\pi}} (-\eta - \varepsilon_i)^{3/2} = \frac{8\pi}{3} \left(\frac{2m_p^*}{h^2} \right)^{3/2} (E_v - F_0)^{3/2}, \quad (3-67)$$

а концентрация электронов находится согласно выражению (3-60).

Уровень Ферми при температуре абсолютного нуля на основании соотношения (3-66) равен:

$$F_0 = E_c + \frac{\hbar^2}{2m_n^*} \left(\frac{3n_0}{8\pi} \right)^{2/3}. \quad (3-68)$$

Из равенств (3-66) и (3-67) следует, что концентрация свободных носителей заряда в полностью вырожденном полупроводнике не зависит от температуры. Она определяется положением уровня Ферми F_0 и величиной 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 (3-68) можно записать следующим образом:

$$\eta = \frac{F - E_c}{k_0 T} = \frac{\hbar^2}{2m_n^* k_0 T} \left(\frac{3n_0}{8\pi} \right)^{2/3} \gg 1. \quad (3-69)$$

Здесь мы заменили η_0 на η , что, как будет показано ниже, для случая сильного вырождения не вносит существенной ошибки. Из анализа условия (3-69) следует, что вырождение наступает тем раньше, чем больше концентрация свободных носителей, чем ниже температура (при данной концентрации) и чем меньше эффективная масса носителей заряда.

Граница перехода от классического случая к вырожденному определяется концентрацией вырождения n_v или температурой вырождения T_v . Поскольку n_v является минимальной концентрацией заряда, соответствующей началу вырождения при данной температуре, поэтому на основании формулы (3-66) можно записать:

$$n_v = \frac{4}{3\sqrt{\pi}} N_c = 3,78 \cdot 10^{15} \left(\frac{m_n^*}{m} \right)^{3/2} T_v^{3/2} = 2 \cdot 10^{19} \left(\frac{m_n^*}{m} \right) \left(\frac{T_v}{300} \right)^{3/2} \quad (3-70)$$

или

$$T_v = 4,1 \cdot 10^{-11} \left(\frac{m}{m_n^*} \right) n_v^{2/3}. \quad (3-70a)$$

Эти выводы подтверждаются экспериментальными данными. На рис. 3-4 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 вырожденным. Область вырождения с увеличением температуры сдвигается к более высоким концентрациям носителей заряда. Если при температуре жидкого азота вырождение наступает при $p_0 \approx 2,5 \cdot 10^{18} \text{ см}^{-3}$, то при комнатной температуре оно имеет место при $p_0 \approx 2 \cdot 10^{19} \text{ см}^{-3}$. Антимонид индия, обладающий очень малой эффективной массой

электронов ($m_n^* = 0,013 m$), при $T = 300^\circ \text{K}$ оказывается вырожденным уже при $n_0 \approx 10^{15} \text{ см}^{-3}$.

Найдем зависимость уровня Ферми для вырожденного полупроводника от температуры. Если $\eta \gg 1$, то интеграл Ферми в выражении (3-43) можно разложить в степенной ряд, и равновесная концентрация электронов n_0 может быть представлена формулой

$$n_0 = N_c \frac{4}{3\sqrt{\pi}} \eta^{3/2} \left(1 + \frac{\pi^2}{8\eta^2} - \frac{7\pi^4}{640\eta^4} + \frac{31\pi^6}{3072\eta^6} - \dots \right). \quad (3-71)$$

Отбросим члены ряда, начиная с третьего, и заменим во втором слагаемом η на η_0 . Можно показать, что при этом несколько снижается ошибка, возникающая при отбрасывании последних членов. В результате для n_0 получим приближенное выражение:

$$n_0 \approx N_c \frac{4}{3\sqrt{\pi}} \eta_0^{3/2} \left(1 + \frac{\pi^2}{8\eta_0^2} \right). \quad (3-71a)$$

Из соотношений (3-66) и (3-71a) имеем:

$$\frac{3\sqrt{\pi} n_0}{4 N_c} = \eta_0^{3/2} \left(1 + \frac{\pi^2}{8\eta_0^2} \right) = \eta_0^{3/2},$$

откуда

$$\eta = \eta_0 \left(1 + \frac{\pi^2}{8\eta_0^2} \right)^{-2/3}. \quad (3-72)$$

При $\eta_0 > 5$ второй член в скобке мал. Разложив выражение (3-72) в ряд и ограничившись двумя членами, найдем окончательно:

$$\eta \approx \eta_0 \left(1 - \frac{\pi^2}{12\eta_0^2} \right) \quad (3-73)$$

или

$$F = F_0 \left[1 - \frac{\pi^2}{12} \left(\frac{k_0 T}{F_0 - E_c} \right)^2 \right]. \quad (3-74)$$

Видно, что при $\eta_0 = (F_0 - E_c) k_0 T \gg 1$ уровень Ферми F в вырожденном полупроводнике практически не зависит от температуры и совпадает с положением уровня F_0 для температуры абсолютного нуля.

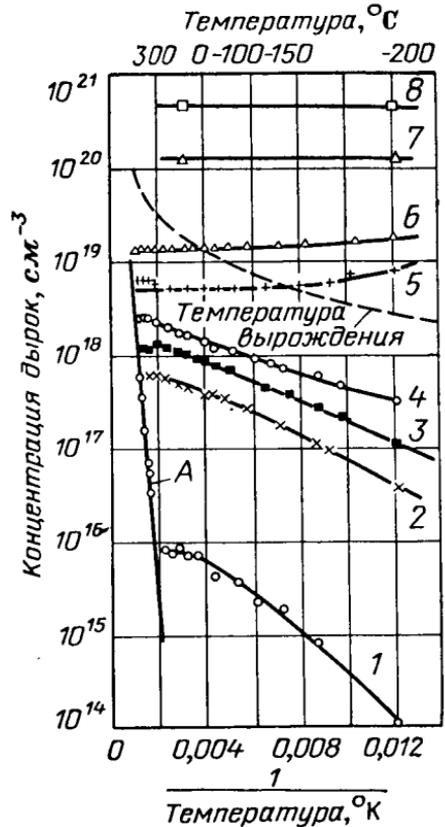


Рис. 3-4. Температурная зависимость концентрации дырок в кремнии, легированном боре.

A — собственная концентрация атомов бора для различных кривых равна (см^{-3}): 1 — нелегированный; 2 — $6,7 \cdot 10^{17}$; 3 — $1,3 \cdot 10^{18}$; 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 — $1,3 \cdot 10^{21}$. Пунктирная кривая показывает связь концентрации дырок с температурой, при которой наступает вырождение.

3-6. СОБСТВЕННЫЙ ПОЛУПРОВОДНИК

Полупроводник будет являться собственным, если влияние примесей для него пренебрежимо мало. Свободные носители возникают только за счет разрыва валентных связей, поэтому в собственном полупроводнике число дырок равно числу свободных электронов, т. е. $n_0 = p_0 = n_i$, что соответствует условию электронейтральности (рис. 3-5). Величина n_i носит название собственной концентрации.

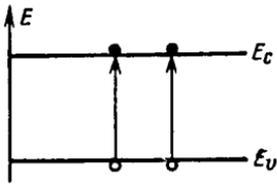


Рис. 3-5. Тепловая генерация носителей заряда в собственном полупроводнике.

Это требование определяет положение уровня Ферми, его можно найти, решив уравнение

$$N_c F_{1/2}(\eta) = N_v F_{1/2}(-\eta - \epsilon_i) \quad (3-75)$$

для каждого конкретного случая.

Невырожденный собственный полупроводник. Для невырожденного полупроводника интеграл Ферми можно аппроксимировать экспонентой (3-59). При этом условии уравнение (3-75) примет вид:

$$N_c e^{\eta} = N_v e^{-\eta - \epsilon_i},$$

откуда легко получить для уровня Ферми

$$F = E_c - \frac{1}{2} \Delta E_g + \frac{k_0 T}{2} \ln \frac{N_v}{N_c} = \frac{E_c + E_v}{2} + \frac{k_0 T}{2} \ln \frac{N_v}{N_c}. \quad (3-76)$$

Если подставить значения N_c из выражения (3-42) и N_v из выражения (3-47), то получим:

$$F = \frac{E_c + E_v}{2} + \frac{3}{4} k_0 T \ln \frac{m_p^*}{m_n^*}. \quad (3-77)$$

Из анализа соотношения (3-77) следует, что при температуре абсолютного нуля уровень Ферми для собственного полупроводника располагается посередине между дном зоны проводимости и вершиной валентной зоны, т. е. $F = \frac{E_c + E_v}{2}$. 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 смещается к дну зоны проводимости, а при $m_p^* < m_n^*$ — к потолку валентной зоны (рис. 3-6, кривые 1, 2 и 3 соответственно).

Определим собственную концентрацию носителей заряда

$$n_i = (n_0 p_0)^{1/2} = (N_c N_v)^{1/2} e^{-\frac{E_c - E_v}{2k_0 T}} = 2 \left(\frac{2\pi k_0}{h^2} \right)^{3/2} (m_n^* m_p^*)^{3/4} T^{3/2} e^{-\frac{\Delta E_g}{2k_0 T}}. \quad (3-78)$$

Таким образом, концентрация собственных носителей n_i зависит от температуры T , ширины запрещенной зоны полупроводника ΔE_g , значений эффективных масс носителей заряда и не зависит от положения уровня Ферми. Температурная зависимость собственной кон-

центрации при $\Delta E_g \gg k_0 T$ определяется в основном экспоненциальным членом уравнения (3-78). Экспериментальные данные температурной зависимости концентрации электрона для германия в области собственной проводимости приведены на рис. 3-7.

Оценим собственную концентрацию носителей заряда в кремнии и германии. Для кремния $m_{dn}^* = 1,08 m$, $m_{dp}^* = 0,59 m$ и при $T = 300^\circ \text{ К}$ $\Delta E_g = 1,08 \text{ эВ}$, а поэтому $n_i \approx 2 \cdot 10^{10} \text{ см}^{-3}$. Для германия $m_{dn}^* = 0,56 m$, $m_{dp}^* = 0,37 m$, $\Delta E_g = 0,66 \text{ эВ}$ и $n_i \approx 2 \cdot 10^{13} \text{ см}^{-3}$.

Если возбуждение носителя заряда происходит за счет теп-

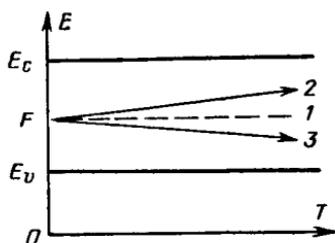


Рис. 3-6. Изменение уровня Ферми с температурой в собственном полупроводнике.

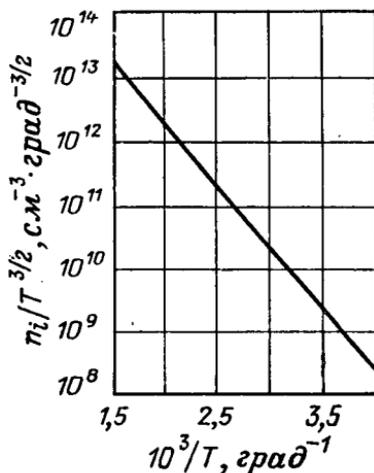


Рис. 3-7. Температурная зависимость $n_i/T^{3/2}$ для германия.

ловой энергии, то вероятность такого процесса пропорциональна $e^{-W/k_0 T}$. Величина W называется энергией активации. Поэтому концентрацию электронов n_i можно выразить и через энергию активации. Она собственно будет равна:

$$n_i = C e^{-\frac{W}{k_0 T}}, \quad (3-79)$$

где C — постоянная величина.

Из сравнения формул (3-78), (3-79) и (3-62) следует, что

$$W = \frac{d \ln(n_i T^{-3/2})}{d \frac{1}{T}} = E_c - F + T \frac{dF}{dT}, \quad (3-80)$$

т. е. энергия активации зависит от положения уровня Ферми.

Применительно к собственному полупроводнику энергия активации

$$W = \frac{1}{2} \Delta E_g, \quad (3-81)$$

т. е. равна половине ширины запрещенной зоны.

Согласно уравнению (3-78) графическая зависимость $\ln(n_i T^{-3/2})$ от $1/T$ должна выражаться прямой линией

$$\ln(n_i T^{-3/2}) = \text{const} - \frac{\Delta E_g}{2k_0} \frac{1}{T}. \quad (3-82)$$

Угол наклона этой прямой $\text{tg } \theta = \frac{\Delta E_g}{2k_0}$. Однако вычисленная по данным рис. 3-7 величина $\Delta E_g = 0,785$ эВ не равна истинной ширине запрещенной зоны германия.

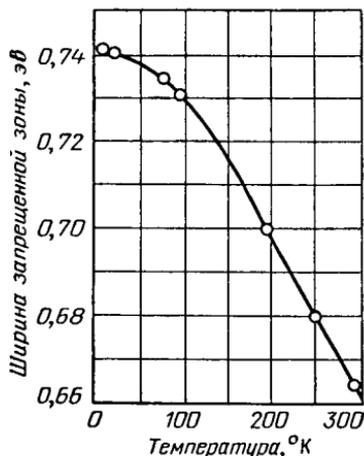


Рис. 3-8. Температурная зависимость ширины запрещенной зоны германия.

Причина этого расхождения заключается в том, что само значение ΔE_g изменяется с температурой. Рост амплитуды тепловых колебаний атомов приводит к ее снижению. Кроме того, с увеличением температуры изменяются межатомные расстояния, что также оказывает влияние на ширину запрещенной зоны. В результате зависимость $\Delta E_g(T)$ может иметь весьма сложный характер. На рис. 3-8 приведена соответствующая кривая для германия.

Из рис. 3-8 видно, что для германия при $T > 200^\circ$ К зависимость $\Delta E_g(T)$ подчиняется линейному закону вида

$$\Delta E_g = \Delta E_{g_0} - \alpha T, \quad (3-83)$$

где ΔE_{g_0} — ширина запрещенной зоны при температуре абсолютного нуля; α — температурный коэффициент.

Подставив выражение (3-83) в уравнение (3-78), получим:

$$\frac{n_i}{T^{3/2}} = 2 \left(\frac{4\pi^2 m_n^* m_p^* k_0^3}{h^4} \right)^{3/4} e^{\frac{\alpha}{2k_0} T} e^{-\frac{\Delta E_{g_0}}{2k_0 T}}, \quad (3-84)$$

откуда найдем, что

$$\ln \frac{n_i}{T^{3/2}} = \text{const} + \frac{\alpha}{2k_0} T - \frac{\Delta E_{g_0}}{2k_0} \frac{1}{T}. \quad (3-85)$$

Таким образом, если ширина запрещенной зоны линейно зависит от температуры, график зависимости $\ln \frac{n_i}{T^{3/2}}$ от $1/T$ также представляет собой прямую линию, наклон которой характеризуется величиной ΔE_{g_0} . Эта величина является шириной запрещенной зоны при 0° К. Истинная ширина запрещенной зоны полупроводника при данной температуре определяется из экспериментально найденной зависимости $\Delta E_g(T)$ (см. рис. 3-8).

Вырожденный собственный полупроводник. У собственного полупроводника согласно соотношению (3-77) скорость изменения уровня Ферми с температурой пропорциональна отношению эффективных масс дырок и электронов. В результате этого с повышением температуры уровень Ферми отдалется от зоны с тяжелыми носителями заряда, приближаясь к зоне с легкими носителями. И если расстояние от уровня Ферми до этой зоны становится соизмеримо с величиной k_0T , то в ней наступает вырождение и соответствующий интеграл Ферми уже не может быть заменен экспонентой. При этом, чем сильнее различаются эффективные массы электронов и дырок, тем раньше наступает вырождение. Если, например, имеем вырождение в зоне проводимости, а в валентной зоне оно отсутствует, то выражение для собственной концентрации носителей заряда будет иметь вид:

$$n_i = N_c F_{1/2}(\eta) = N_c e^{-\eta - \epsilon_i}. \quad (3-86)$$

Отсюда получается следующая формула, определяющая положение уровня Ферми:

$$\epsilon^n F_{1/2}(\eta) = \left(\frac{m_p^*}{m_n^*} \right)^{3/2} e^{-\epsilon_i}. \quad (3-87)$$

Классическим примером полупроводника с сильно различающимися эффективными массами является антимонид индия, для которого $m_p^* \approx 10 m_n^*$. На рис. 3-9 показано изменение положения уровня Ферми в InSb в зависимости от температуры. При этом учитывается зависимость $\Delta E_g(T)$, которая согласно экспериментальным данным имеет вид:

$$\Delta E_g = (0,26 - 2,7 \cdot 10^{-4} T), \text{ эв.}$$

Как видно из зависимости, приведенной на рис. 3-9, уже при $T = 200^\circ \text{K}$ уровень Ферми лежит ниже E_c всего на $4 k_0 T$, а около 450°K пересекает уровень E_c и входит в зону проводимости, что приводит к сильному вырождению электронного газа. Это необходимо учитывать, и при расчете величины n_i следует использовать равенство (3-86).

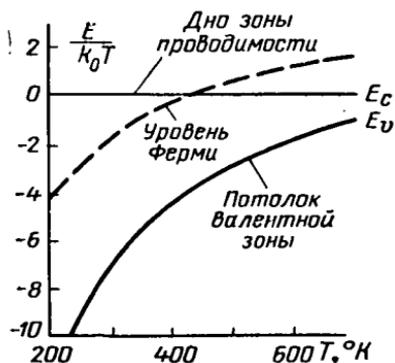


Рис. 3-9. Положение уровня Ферми (в единицах k_0T) в зависимости от температуры для антимонида индия.

3-7. ЗАВИСИМОСТЬ ПОЛОЖЕНИЯ УРОВНЯ ФЕРМИ ОТ КОНЦЕНТРАЦИИ ПРИМЕСИ И ТЕМПЕРАТУРЫ ДЛЯ НЕВЫРОЖДЕННОГО ПОЛУПРОВОДНИКА

Полупроводник может содержать как донорную, так и акцепторную примесь. Пусть концентрация этих примесей соответственно

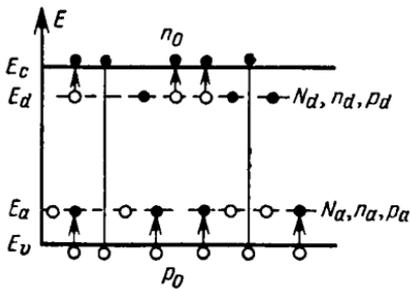


Рис. 3-10. Полупроводник, имеющий примеси донорного и акцепторного типа.

равна N_d и N_a . Энергетическая схема такого полупроводника представлена на рис. 3-10. Электроны в валентной зоне возникают в результате теплового возбуждения электронов из валентной зоны и атомов примеси. Эти переходы показаны стрелками.

Для определения положения уровня Ферми в полупроводнике используется условие электронейтральности, по которому в кристалле суммарный заряд всех заряженных частиц должен быть равен нулю. Для полупроводника с двумя видами примеси уравнение электронейтральности запишется в виде

$$n_0 + n_a = p_0 + p_d, \quad (3-88)$$

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 этих выражений для n_a и p_d условие электронейтральности (3-88) можно представить в виде

$$(n_0 + n_d) - (p_0 + p_a) = N_d - N_a. \quad (3-89)$$

Выразив концентрации носителей заряда в соответствии с формулами (3-41), (3-46), (3-49) и (3-51), получим:

$$N_c F_{1/2}(\eta) + \frac{N_d}{\beta^{-1} e^{-\epsilon_d - \eta} + 1} - N_v F_{1/2}(-\eta - \epsilon_i) - \frac{N_a}{\beta^{-1} e^{\epsilon_a + \eta} + 1} = N_d - N_a. \quad (3-90)$$

Для невырожденного полупроводника равенство (3-90) значительно упрощается

$$N_c e^{\eta} + \frac{N_d}{\beta^{-1} e^{-\eta - \epsilon_d} + 1} - N_v e^{-\eta - \epsilon_i} - \frac{N_a}{\beta^{-1} e^{\eta + \epsilon_a} + 1} = N_d - N_a. \quad (3-91)$$

Однако и в этом случае определение уровня Ферми встречает большие трудности. Поэтому в дальнейшем мы 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 (3-89) запишется при этом так:

$$n_0 + n_d = N_d \quad (3-92)$$

или

$$n_0 = P_d,$$

откуда согласно равенству (3-50) будем иметь:

$$n_0 = \frac{N_d}{\beta e^{\eta + \varepsilon_d} + 1}. \quad (3-93)$$

Считаем, что вырождение отсутствует. Следовательно, $n_0 = N_c e^{\eta}$. Из этого равенства найдем, что

$$e^{\eta} = \frac{n_0}{N_c}. \quad (3-94)$$

Подставим (3-94) в формулу (3-93), в результате получим уравнение для концентрации электронов в зоне проводимости

$$n_0 = \frac{N_c N_d}{N_c + \beta e^{\varepsilon_d} n_0}, \quad (3-95)$$

которое легко преобразуется к виду

$$n_0^2 + \beta^{-1} e^{-\varepsilon_d} N_c n_0 - \beta^{-1} e^{-\varepsilon_d} N_d N_c = 0. \quad (3-96)$$

Решая полученное квадратное уравнение относительно n_0 , находим:

$$n_0 = \frac{\beta^{-1} e^{-\varepsilon_d} N_c}{2} \left(\sqrt{1 + 4\beta e^{\varepsilon_d} N_c^{-1} N_d} - 1 \right). \quad (3-97)$$

Минус перед корнем отброшен, так как подкоренное выражение больше единицы, а $n_0 > 0$. Это соотношение можно свести к следующему выражению:

$$n_0 = \frac{2N_d}{\sqrt{1 + 4\beta e^{\varepsilon_d} N_c^{-1} N_d} + 1}. \quad (3-98)$$

Положение уровня Ферми определится из соотношения (3-94)

$$e^{\eta} = \frac{n_0}{N_c} = \frac{2N_d/N_c}{\sqrt{1 + 4\beta e^{\varepsilon_d} N_c^{-1} N_d} + 1}. \quad (3-99)$$

Прологарифмировав равенство (3-99), получим:

$$\eta = -\ln \left[\frac{N_c}{2N_d} + \sqrt{\left(\frac{N_c}{2N_d} \right)^2 + \beta e^{\varepsilon_d} N_c N_d^{-1}} \right] \quad (3-100)$$

или

$$F = E_c - k_0 T \ln \left[\frac{N_c}{2N_d} + \sqrt{\left(\frac{N_c}{2N_d} \right)^2 + \beta e^{\frac{E_c - E_d}{k_0 T}} N_c N_d^{-1}} \right]. \quad (3-101)$$

Проведем анализ найденного выражения для уровня Ферми. В области низких температур, где справедливо условие $e^{\frac{E_c - E_d}{k_0 T}} \gg \frac{N_c}{N_d}$, формула (3-101) упрощается и положение уровня Ферми будет определяться соотношением вида

$$F = E_c - \frac{E_c - E_d}{2} - \frac{k_0 T}{2} \ln \frac{\beta N_c}{N_d}, \quad (3-102)$$

а энергия активации согласно формуле (3-80) будет:

$$W = \frac{E_c - E_d}{2},$$

т. е. равна половине энергии ионизации донорной примеси.

Как следует из соотношений (3-102), в донорном полупроводнике при температуре абсолютного нуля уровень Ферми располагается посередине между дном зоны проводимости и уровнем донорной примеси. В области достаточно низких температур (порядка нескольких градусов Кельвина), когда $2N_c < N_d$, уровень Ферми вначале повышается до некоторого максимального значения, а затем начинает снижаться, и при $2N_c = N_d$ снова имеем $F = \frac{E_c + E_d}{2}$, как и для случая $T = 0$. Дальнейшее повышение температуры сопровождается ростом N_c , и в области, когда $2N_c > N_d$, уровень Ферми продолжает снижаться. Такому перемещению уровня Ферми соответствует экспоненциальная температурная зависимость концентрации электронов

$$n_0 = N_c e^n = N_c e^{\left(-\frac{E_c - E_d}{2k_0 T} - \frac{1}{2} \ln \frac{N_c}{\beta^{-1} N_d} \right)} = (\beta^{-1} N_c N_d)^{1/2} e^{-\frac{E_c - E_d}{2k_0 T}}. \quad (3-103)$$

Эта область изменения уровня Ферми с температурой, которая описывается формулой (3-102), является областью слабой ионизации примеси. Она обозначена цифрой 1 на рис. 3-11, на котором проиллюстрировано изменение уровня Ферми (рис. 3-11, а) и концентрации электронов (рис. 3-11, б) с температурой для донорного полупроводника.

При дальнейшем повышении температуры концентрация электронов в зоне проводимости становится сравнимой с концентрацией примеси N_d , и выражения (3-102) и (3-103) неприменимы. Однако теперь можно рассматривать другой крайний случай, когда температура достаточно высокая и выполняется неравенство $e^{(E_c - E_d)/k_0 T} \ll \beta N_c / N_d$. При этом формула (3-101) аппроксимируется выражением

$$F = E_c - k_0 T \ln \left(\frac{N_c}{N_d} \right), \quad (3-104)$$

которому соответствует

$$n_0 = N_d. \quad (3-105)$$

Это означает, что практически вся донорная примесь ионизирована, и концентрация электронов в зоне проводимости не зависит от температуры. Эта область температур, при которой имеет место полная ионизация примеси, носит название области истощения примеси (или области насыщения) и на рис. 3-11 отмечена цифрой 2.

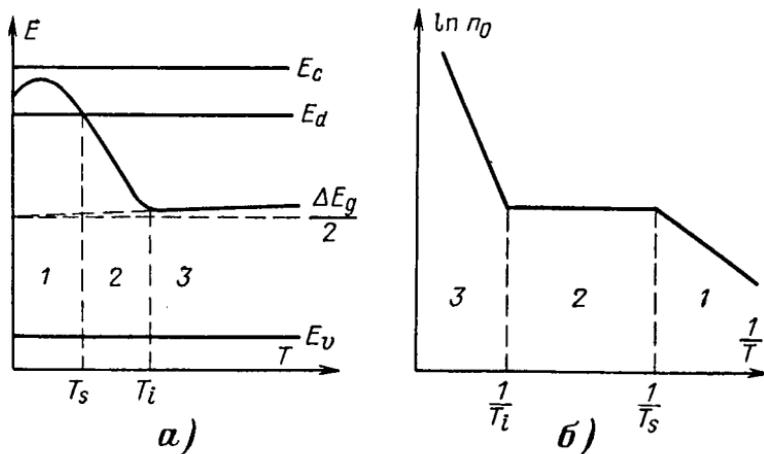


Рис. 3-11. Изменение положения уровня Ферми (а) и концентрации электронов (б) с температурой для донорного полупроводника.

Условие полной ионизации донорной примеси, когда $n_0 = N_d$, соответствует положению уровня Ферми на несколько k_0T 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 насыщения T_s , и ее можно определить из условия

$$F = E_c - k_0T_s \ln \left(\frac{N_c}{N_d} \right) = E_d. \quad (3-106)$$

Отсюда

$$T_s = \frac{E_c - E_d}{k_0 \ln(N_c/N_d)}. \quad (3-107)$$

Как следует из формулы (3-107), температура насыщения тем ниже, чем меньше энергия ионизации $E_c - E_d$ и концентрация N_d донорной примеси и чем больше эффективная масса электронов, определяющая величину N_c . При малых значениях $E_c - E_d$ насыщение наступает при очень низких температурах. Например, в электронном германии, легированном сурьмой в количестве

$N_d \approx 10^{16} \text{ см}^{-3}$, для которой энергия ионизации равна 0,0096 эв, насыщение наступает уже при 32° 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 (3-76), (3-106) и (3-78). На рис. 3-11 область 3 соответствует области собственной проводимости. В этом случае $F = F_i$. Из (3-76) и (3-106) следует, что

$$F_i = E_c - \frac{\Delta E_g}{2} - \frac{k_0 T_i}{2} \ln \frac{N_c}{N_v} = E_c - k_0 T_i \ln \frac{N_c}{N_d}, \quad (3-108)$$

откуда получаем:

$$T_i = \frac{\Delta E_g}{k_0 \ln \left(\frac{N_c N_v}{N_d^2} \right)}. \quad (3-109)$$

Анализ этого выражения показывает, что температура T_i , при которой наступает собственная проводимость у донорного полу-

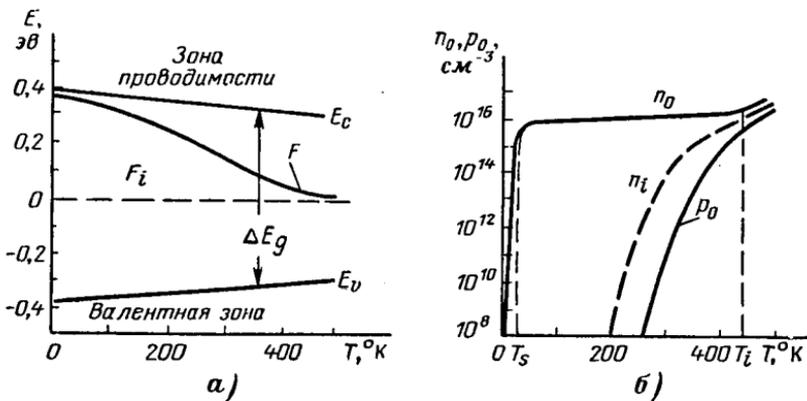


Рис. 3-12. Температурная зависимость уровня Ферми (а) и концентрации (б) носителей заряда для германия, легированного сурьмой.

проводника, 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. 3-12, а приведены температурные зависимости уровня Ферми и на рис. 3-12, б — концентрация равновесных электронов n_0 и дырок p_0 для германия, легированного сурьмой в количестве $N_d = 10^{16} \text{ см}^{-3}$. Кроме того, на этих кривых

пунктиром показан ход F_i и n_i в собственном полупроводнике. При построении графиков учтена зависимость ширины запрещенной зоны германия от температуры. Из этого рисунка видно, 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 незначительно поднимается к E_c , проходит максимум и начинает опускаться. Понижаясь, уровень Ферми приближается к уровню примеси E_d . При температуре насыщения T_s на донорной примеси находится электронов в количестве, равном

$$n_d = \frac{N_d}{\frac{E_d - F}{k_0 T} + 1} = \frac{2}{3} N_d,$$

а в зоне проводимости соответственно будет $1/3 N_d$ электронов. С дальнейшим ростом температуры уровень Ферми продолжает опускаться и наступает область истощения; вся примесь ионизирована 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

$$p_0 = \frac{n_i^2}{n_0} = \frac{N_c N_v}{N_d} e^{-\frac{\Delta E_g}{k_0 T}}. \quad (3-110)$$

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 ($N_d = 0$)

Решение задачи для полупроводника, легированного акцепторной примесью, полностью аналогично случаю донорного полупроводника. Условие электронейтральности теперь будет иметь вид:

$$p_0 = n_a = N_a - p_a. \quad (3-111)$$

Отсюда для невырожденного полупроводника с учетом формул (3-61) и (3-52) легко получить соотношение для p_0 , аналогичное (3-93) и (3-95):

$$p_0 = \frac{N_a}{\beta e^{-\eta - \epsilon_a} + 1} = \frac{N_a}{N_v + \beta e^{e_i - \epsilon_a} p_0}. \quad (3-112)$$

После несложных алгебраических вычислений находим:

$$p_0 = \frac{2N_a}{\sqrt{1 + 4\beta e^{e_i - \epsilon_a} N_a N_v^{-1}} + 1}. \quad (3-113)$$

Положение уровня Ферми будет определяться равенством

$$e^{n + \varepsilon_i} = \frac{N_v}{p_0} = \frac{N_v}{2N_a} \left[\sqrt{1 + 4\beta e^{\varepsilon_i - \varepsilon_a} N_a N_v^{-1}} + 1 \right]. \quad (3-114)$$

После логарифмирования получим выражение для уровня Ферми

$$F = E_c - \Delta E_g + k_0 T \ln \left[\frac{N_v}{2N_a} + \sqrt{\left(\frac{N_v}{2N_a} \right)^2 + \beta \frac{N_v}{N_a} e^{\frac{E_a - E_v}{k_0 T}}} \right]. \quad (3-115)$$

В области низких температур, когда выполняется условие $e^{\frac{E_a - E_v}{k_0 T}} \gg \frac{N_v}{N_a}$, имеем:

$$\begin{aligned} F &= E_c - \Delta E_g + \frac{1}{2}(E_a - E_v) + \frac{k_0 T}{2} \ln \frac{\beta N_v}{N_a} = \\ &= \frac{E_v + E_a}{2} + \frac{k_0 T}{2} \ln \frac{\beta N_v}{N_a}. \end{aligned} \quad (3-116)$$

В дырочном полупроводнике при температуре абсолютного нуля уровень Ферми лежит посередине между потолком валентной зоны и уровнем акцепторной примеси, т. е. $F = (E_v + E_a)/2$. С увеличением температуры уровень Ферми вначале опускается, когда $N_a > \beta N_v$, а затем поднимается до уровня E_a . Концентрация свободных дырок в этой области определяется соотношением

$$p_0 = N_v e^{-n - \varepsilon_i} = (\beta^{-1} N_a N_v)^{1/2} e^{\frac{E_v - E_a}{2k_0 T}}. \quad (3-117)$$

Положение уровня Ферми в зависимости от температуры для акцепторного полупроводника изображено на рис. 3-13.

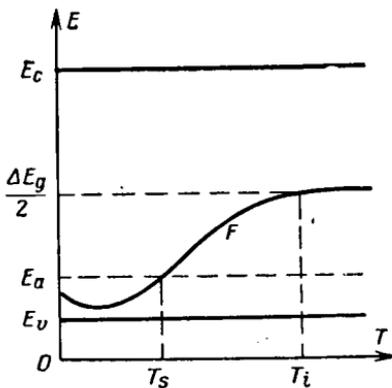


Рис. 3-13. Изменение положения уровня Ферми с температурой для акцепторного полупроводника.

При высоких температурах, когда $e^{(E_a - E_v)/k_0 T} \ll \frac{N_v}{N_a}$, и из выражения (3-115) получаем:

$$F = E_c - \Delta E_g + k_0 T \ln \frac{N_v}{N_a}, \quad (3-118)$$

а концентрация свободных дырок равна:

$$p_0 = N_v e^{-\ln N_v / N_a} = N_a. \quad (3-119)$$

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.

3-8. 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 акцепторных центров согласно схеме, изображенной на рис. 3-10.

Рассмотрим сначала случай, когда в полупроводник введены в равных количествах доноры и акцепторы, 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 называется **с к о м п е н с и р о в а н н ы м**. В нем имеет место взаимная компенсация примесей, а поэтому они не могут являться поставщиками электронов и дырок. С ростом температуры свободные носители заряда будут возникать только благодаря ионизации атомов основного вещества и полупроводник ведет себя, по-существу, как собственный. Однако в скомпенсированном 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 величину $N_d - N_a$. Пока не наступит собственная проводимость, свободные дырки в валентной зоне отсутствуют и полупроводник ведет себя как чисто донорный. Такой полупроводник называется **ч а с т и ч н о к о м п е н с и р о в а н н ы м**. Степень компенсации определяется отношением N_a/N_d . В переносе заряда принимают участие только электроны, и постоянная Холла в таком материале имеет отрицательный знак.

В случае, когда $N_a > N_d$ и уровень Ферми проходит значительно ниже E_d , донорная примесь полностью ионизирована. 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 ионам 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.

Найдем зависимость концентрации свободных электронов и уровня Ферми от температуры в полупроводнике с частично компенсированной примесью для случая $N_d > N_a$. Условие электро neutrality при этом имеет вид:

$$n_0 + n_d = N_d - \bar{N}_a \quad (3-120)$$

или

$$n_0 + N_a = p_d. \quad (3-121)$$

Будем считать, что полупроводник невырожден. Используя соотношение $n_0 = N_c e^{\eta}$, получаем из формул (3-121) и (3-50)

$$n_0 + N_a = \frac{N_c N_d}{N_c + \beta e^{e_d} n_0}. \quad (3-122)$$

Решив это уравнение, найдем:

$$n_0 = \frac{2(N_d - N_a)}{1 + \frac{\beta N_a}{N_c} e^{e_d} + \left[\left(1 + \frac{\beta N_a}{N_c} e^{e_d} \right)^2 + \frac{4(N_d - N_a) e^{e_d}}{\beta^{-1} N_c} \right]^{1/2}}. \quad (3-123)$$

Из соотношения $e^{\eta} = n_0 / N_c$ определим уровень Ферми, который будет равен:

$$F = E_c + k_0 T \times \ln \frac{2 \frac{N_d - N_a}{N_c}}{1 + \frac{\beta N_a}{N_c} e^{\frac{E_c - E_d}{k_0 T}} + \left\{ \left(1 + \frac{\beta N_a}{N_c} e^{\frac{E_c - E_d}{k_0 T}} \right)^2 + \frac{4\beta \left[(N_d - N_a) e^{\frac{E_c - E_d}{k_0 T}} \right]}{N_c} \right\}^{1/2}}}. \quad (3-124)$$

Нетрудно видеть, что при $N_a = 0$ найденные выражения совпадают с формулами (3-98) и (3-101) для невырожденного донорного полупроводника.

В области высоких температур, когда донорные атомы полностью ионизированы, справедливы неравенства:

$$\frac{\beta N_a}{N_c} e^{e_d} \ll 1 \quad \text{и} \quad \frac{4\beta (N_d - N_a) e^{e_d}}{N_c} \ll 1. \quad (3-125)$$

Учет этих неравенств дает для n_0 и F следующие соотношения:

$$n_0 = N_d - N_a; \quad (3-126)$$

$$F = E_c - k_0 T \ln \frac{N_d - N_a}{N_c}. \quad (3-127)$$

Как и следовало ожидать, концентрация свободных электронов постоянна и определяется числом нескомпенсированных донорных атомов. Уровень Ферми при $T = 0$ совпадает с E_c , а с ростом температуры понижается, так как согласно формуле (3-42) $N_c \sim T^{3/2}$. При температурах, соответствующих началу собственной проводимости, он 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

$$\frac{\beta N_a}{N_c} e^{\epsilon_d} \ll 1 \ll \frac{4\beta (N_d - N_a) e^{\epsilon_d}}{N_c}. \quad (3-128)$$

В этом интервале температур из соотношений (3-123) и (3-124) получаем приближенные формулы

$$n_0 = (\beta^{-1} N_c N_d)^{1/2} e^{-\frac{E_c - E_d}{2k_0 T}}; \quad (3-129)$$

$$F = \frac{E_c + E_d}{2} + \frac{k_0 T}{2} \ln \frac{N_d}{\beta N_c}, \quad (3-130)$$

совпадающие с выражениями (3-103) и (3-102) для донорного полупроводника. Понятно, что при $n_0 \gg N_a$ концентрация свободных электронов не зависит от количества акцепторных атомов, как и следует из равенства (3-129). При $T = 0$ уровень Ферми лежит в середине между E_c и E_d . С ростом температуры он понижается (кривая 1, рис. 3-14, а). Однако при более низких температурах, когда $n_0 \ll N_a$, влиянием акцепторов уже нельзя пренебречь. Положение уровня Ферми в этой области [выражение (3-124)] существенно отличается от того, что имеет место в чисто донорном полупроводнике [соотношение (3-102)]. Согласно выражению (3-124) при $T = 0$ он всегда совпадает с дном зоны проводимости E_c .

В случае сильной компенсации справедливо неравенство $n_0 \ll N_a < N_d$, и из выражений (3-123) и (3-124) следует:

$$n_0 = \beta^{-1} N_c \frac{N_d - N_a}{N_a} e^{-\frac{E_c - E_d}{k_0 T}}; \quad (3-131)$$

$$F = E_d + k_0 T \ln \frac{N_d - N_a}{\beta N_a} = E_d + k_0 T \ln \frac{N_d - N_a}{2N_a}. \quad (3-132)$$

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.

Ход зависимостей $F(T)$ и $n_0(T)$ для различной степени компенсации показан на рис. 3-14.

При повышении температуры начиная от абсолютного нуля уровень Ферми повышается от уровня донорной примеси, если $(N_d - N_a)/2N_a > 1$ (кривая 2). В противоположном случае, когда

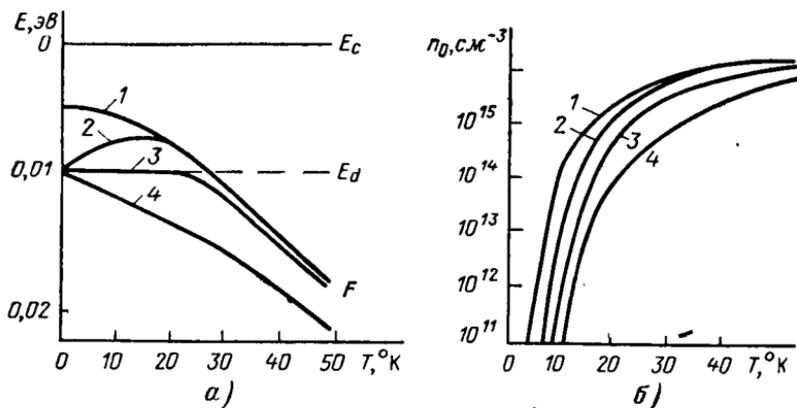


Рис. 3-14. Зависимость положения уровня Ферми (а) и концентрации электронов (б) от температуры при различной степени компенсации.

$(N_d - N_a)/2N_a < 1$, уровень Ферми с ростом температуры снижается (кривая 4), а при $(N_d - N_a)/2N_a = 1$ он остается неизменным, пока выполняется условие $n_0 \ll N_a$ (кривая 3).

Из сравнения кривых рис. 3-14, б видно, что температура, при которой наступает насыщение зависимости $n_0(T)$, растет с увеличением степени компенсации донорной примеси.

3-9. ПРИМЕСНЫЕ ЗОНЫ

Отметим еще одну особенность примесных полупроводников. У донорного полупроводника при температурах, при которых еще нет 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.

$$n_0 = p_d = \frac{N_d}{\frac{F - E_d}{2e^{k_0 T}} + 1}. \quad (3-133)$$

В таком полупроводнике при температуре, приближающейся к 0°K , концентрация электронов n_0 должна уменьшаться до нуля,

а постоянная Холла $R = -\frac{1}{en_0}$ должна монотонно возрастать.

Однако результаты исследования материалов с высокой концентрацией примесных атомов показывают, что при снижении температуры значение n_0 стремится к конечной величине, а постоянная Холла при некоторой температуре проходит через максимум, как это видно из данных рис. 3-15 и 3-16, где изображены кривые удельного сопротивления и коэффициента Холла дырочных образцов

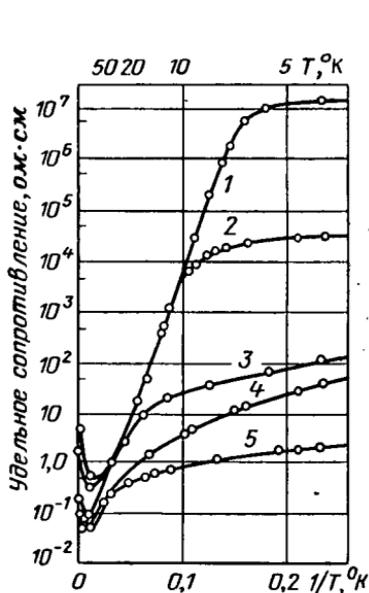


Рис. 3-15: Удельное сопротивление дырочных образцов германия как функция температуры при низких ее значениях.

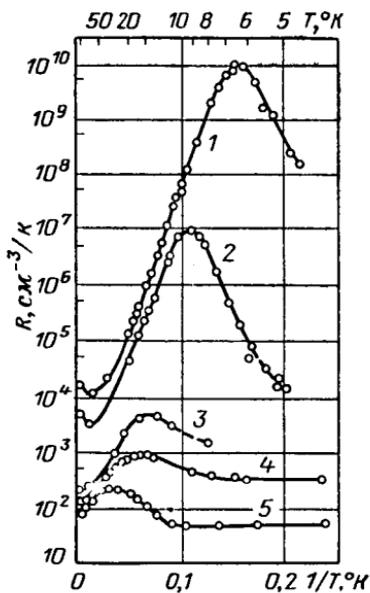


Рис. 3-16: Коэффициент Холла дырочных образцов германия как функция температуры при низких ее значениях.

германия как функции температуры. Образцы имеют разную степень легирования, которая увеличивается с ростом номера кривой.

Кроме того, у полупроводников при увеличении концентрации примеси имеет место снижение энергии ионизации примесных центров. При достаточно больших концентрациях атомов примеси энергия ионизации их обращается в нуль. В германии и кремнии, легированных донорами 5 группы, это происходит при $3 \cdot 10^{17} \text{ см}^{-3}$ и 10^{18} см^{-3} соответственно. В качестве примера для кремния, легированного бором, эта зависимость представлена на рис. 3-17.

Из этих экспериментальных фактов следует, что в полупроводнике, содержащем примеси, при низких температурах имеет место особое явление, получившее название примесной проводимости. При этом механизм примесной проводимости раз-

личен в зависимости от концентрации введенной примеси. Для примесной проводимости характерна исключительно сильная зависимость от концентраций примеси. Как следует из рис. 3-15, при изменении концентрации примеси в германии в 30 раз примесная проводимость меняется в 10^7 раз.

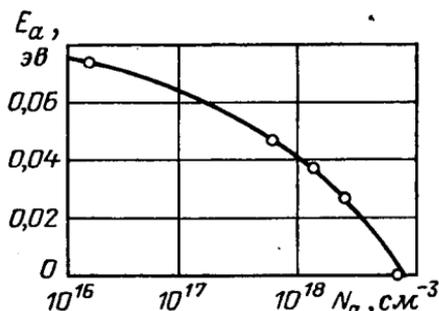


Рис. 3-17. Зависимость энергии ионизации атомов бора от его концентрации в кремнии.

трической проницаемости и обратно пропорционален эффективной массе электрона. В силу этого эффективный радиус орбиты для электрона примеси в германии в 60 раз больше борковского радиуса орбиты электрона в атоме водорода. Поэтому уже при концентрации примеси примерно 10^{15} см^{-3} волновые функции электронов соседних примесных атомов заметно перекрываются.

В реальном донорном полупроводнике всегда имеется некоторое количество акцепторных центров N_a и часть электронов с донорных уровней N_d переходит на уровни акцепторов. Оставшиеся $N_d - N_a$ электроны будут распределяться между зоной проводимости и донорными уровнями. При температуре абсолютного нуля все электроны будут находиться на донорных уровнях. Если имеет место перекрытие волновых функций электронов соседних донорных центров,

то при достаточно низких температурах электрон посредством туннельного эффекта будет перемещаться от одного центра к другому. Так как примесные центры расположены на больших расстояниях друг от друга, то подвижность носителей, перемещающихся по примесным уровням, значительно меньше подвижности электронов в зоне проводимости. В области же высоких температур будет

Водородоподобная модель примесных атомов не учитывает взаимодействия их между собой. Однако при достаточно высоких концентрациях атомов примеси учет их взаимодействия необходим, поскольку борковский радиус для электрона, связанного на водородоподобном примесном центре, согласно выражению (2-136) пропорционален диэлек-

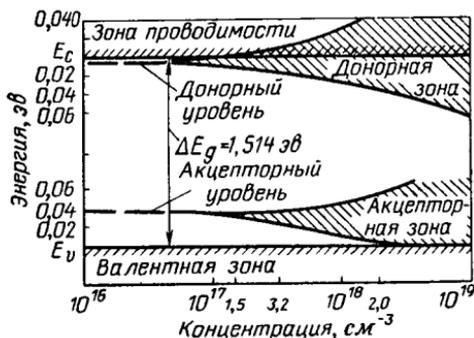


Рис. 3-18. Появление донорной и акцепторной зон в арсениде галлия и их слияние с зоной проводимости и валентной при увеличении концентрации примеси.

преобладать обычная проводимость по зонам энергии, поэтому для всей области изменения температуры проводимость σ и коэффициент Холла R такого образца будут определяться выражениями

$$\sigma = en_1\mu_1 + en_2\mu_2; \quad (3-134)$$

$$R = \frac{n_1\mu_1\mu_{1H} + n_2\mu_2\mu_{2H}}{e(n_1\mu_1 + n_2\mu_2)}, \quad (3-135)$$

где n_1, μ_1, μ_{1H} и n_2, μ_2, μ_{2H} — концентрации, дрейфовые и холловские подвижности электронов в зоне проводимости и для примесной проводимости соответственно. Если пренебречь зависимостью подвижностей от температуры, постоянная Холла имеет максимальное значение при

$$n_1\mu_1 = n_2\mu_2. \quad (3-136)$$

При критической концентрации примеси, равной N_c , перекрытие электронных волновых функций между центрами становится достаточно большим и примесный уровень превращается в зону, расположенную вблизи дна зоны проводимости. В этом случае возникает другая форма примесной проводимости, обусловленная электронами, которые ведут себя подобно вырожденному электронному газу. При больших концентрациях примеси примесная зона, расширяясь, может перекрываться с зоной проводимости (или с валентной зоной в случае акцепторного полупроводника). В этом случае вырождение наблюдается в широком интервале температур.

На рис. 3-18 приведено изменение донорной и акцепторной зон в арсениде галлия при увеличении концентрации примеси.

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^{20} см^{-3} перекрытие зон велико, 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 качественно на рис. 3-19.

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.

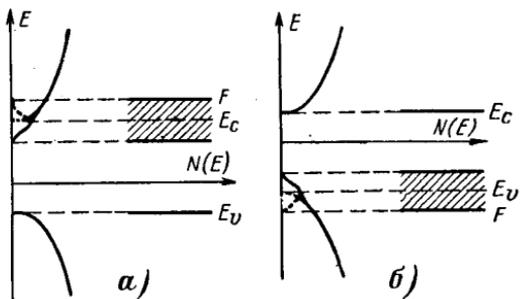


Рис. 3-19. Плотность состояний и положение уровня Ферми в вырожденном электронном (а) и дырочном (б) полупроводниках.

МЕХАНИЗМЫ РАССЕЯНИЯ НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКАХ

4-1. КИНЕТИЧЕСКОЕ УРАВНЕНИЕ БОЛЬЦМАНА

До сих пор мы рассматривали совокупность носителей заряда в кристалле как систему частиц, находящихся в термодинамическом равновесии. В этом случае поведение носителей описывается равновесной функцией распределения (3-27), зависящей в общем случае от энергии частиц.

Если наложить на кристалл внешнее поле (электрическое, магнитное, температурное), то движение носителей заряда в кристалле приобретает упорядоченный направленный характер. Явления, происходящие при этом в кристалле, носят название явления переноса или кинетических явлений. Естественно, что система носителей теперь будет описываться другой, неравновесной функцией распределения, зависящей от энергии частиц, их координат и времени $f(\mathbf{r}, \mathbf{k}, t)$.

Для того чтобы определить концентрацию электронов в неравновесном состоянии, поступим следующим образом. В фазовом пространстве кристалла единичного объема выделим элемент объема $d\Gamma$. Он определяется объемом обычного пространства $d\tau_r = dx dy dz$ и объемом пространства квазиимпульсов $d\tau_p = \hbar^3 d\tau_k$, тогда

$$d\Gamma = d\tau_r d\tau_p = \hbar^3 d\tau_r d\tau_k. \quad (4-1)$$

В элементе объема фазового пространства имеется $d\Gamma/\hbar^3$ фазовых ячеек, в каждой из которых могут быть два электрона с противоположно направленными спинами. Следовательно, в элементе объема фазового пространства содержится $2 \frac{d\Gamma}{\hbar^3}$ квантовых состояний. Если $f(\mathbf{r}, \mathbf{k}, t)$ есть вероятность нахождения электрона в этих состояниях, то число электронов в элементе фазового пространства в момент времени t будет равно:

$$dn = f(\mathbf{k}, \mathbf{r}, t) 2 \frac{d\Gamma}{\hbar^3} = f(\mathbf{r}, \mathbf{k}, t) \frac{d\tau_k}{4\pi^3} d\tau_r. \quad (4-2)$$

Выведем уравнение, которому удовлетворяет неравновесная функция распределения $f(\mathbf{r}, \mathbf{k}, t)$. Для этого рассмотрим движение электронов в обычном пространстве и в пространстве волнового вектора.

Сначала рассмотрим систему электронов в обычном, геометрическом пространстве. Пусть внешнее силовое поле приложено таким образом, что электроны движутся со скоростью v_x вдоль положительного направления оси x . Рассчитаем изменение числа электронов за время dt внутри элемента объема $d\tau_r$.

В элемент объема $d\tau_r$, который изображен на рис. 4-1, через его левую грань за время dt согласно (4-2) войдет электронов,

движущихся со скоростью v_x , в количестве, равном

$$f(\kappa, x, y, z, t) \frac{d\tau_\kappa}{4\pi^3} v_x dy dz dt. \quad (4-3)$$

За это же время число электронов, ушедших из объема через правую грань, составит:

$$-f(\kappa, x + dx, y, z, t) \frac{d\tau_\kappa}{4\pi^3} v_x dy dz dt. \quad (4-4)$$

Таким образом, за время dt количество электронов в элементе объема $d\tau_r$ уменьшится на величину

$$\begin{aligned} [f(\kappa, x, y, z, t) - f(\kappa, x + dx, y, z, t)] \frac{d\tau_\kappa}{4\pi^3} v_x dy dz dt = \\ = -v_x \frac{\partial f}{\partial x} \frac{d\tau_\kappa}{4\pi^3} dx dy dz dt = -v_x \frac{\partial f}{\partial x} \frac{d\tau_\kappa}{4\pi^3} d\tau_r dt. \end{aligned} \quad (4-5)$$

В общем случае произвольного направления движения носителей заряда со скоростью \mathbf{v} с составляющими v_x, v_y, v_z изменение количества электронов с данным κ в элементе объема $d\tau_r$ за время dt будет равно:

$$\begin{aligned} - \left[v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y} + v_z \frac{\partial f}{\partial z} \right] \frac{d\tau_\kappa}{4\pi^3} d\tau_r dt = \\ = -(\mathbf{v} \nabla_r f) \frac{d\tau_\kappa}{4\pi^3} d\tau_r dt. \end{aligned} \quad (4-6)$$

Это изменение количества электронов обусловлено процессами переноса, происходящими под влиянием диффузии носителей заряда вдоль градиентов температуры или их концентрации из-за неоднородности вещества.

Рассмотрим теперь поведение электронов в пространстве волнового вектора. Подсчитаем изменение числа электронов за время dt за счет прихода и ухода их через соответствующие грани элемента объема $d\tau_\kappa$. Для этого нужно, так же как и в случае движения электронов в обычном пространстве, подсчитать количество пришедших и ушедших электронов через все грани элемента объема $d\tau_\kappa$. В результате подсчета получим, что изменение количества электронов за время dt при движении их в пространстве волнового вектора составляет величину, равную

$$\begin{aligned} - \left(\frac{\partial \kappa_x}{\partial t} \frac{\partial f}{\partial \kappa_x} + \frac{\partial \kappa_y}{\partial t} \frac{\partial f}{\partial \kappa_y} + \frac{\partial \kappa_z}{\partial t} \frac{\partial f}{\partial \kappa_z} \right) \frac{d\tau_\kappa}{4\pi^3} d\tau_r dt = - \left(\frac{d\kappa}{dt} \nabla_\kappa f \right) \frac{d\tau_\kappa}{4\pi^3} d\tau_r dt = \\ = - \frac{1}{\hbar} (\mathbf{F} \nabla_\kappa f) \frac{d\tau_\kappa}{4\pi^3} d\tau_r dt. \end{aligned} \quad (4-7)$$

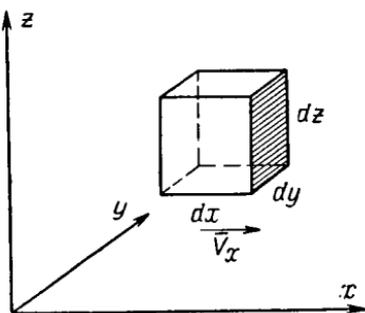


Рис. 4-1. Элементарный объем в геометрическом пространстве.

Здесь

$$\hbar \frac{d\mathbf{k}}{dt} = \frac{d\mathbf{p}}{dt} = \mathbf{F}(\mathbf{r}, t) \quad (4-8)$$

есть сила, действующая на электрон в точке \mathbf{r} в момент времени t . Соотношение (4-7) выражает собой изменение количества электронов в результате действия внешней силы, обусловленной электрическими и магнитными 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 порядка 10^{-7} см, которое охватывает несколько периодов решетки. Так как тепловая скорость электрона порядка 10^7 см/сек, то время взаимодействия со структурными нарушениями составляет величину 10^{-14} сек. Такое кратковременное взаимодействие не может привести к заметному изменению координаты, но сильно изменит скорость и квазимпульс электрона. Этот процесс аналогичен удару в механике, поэтому он и получил название столкновения или соударения. При соударении происходит изменение числа электронов, движущихся в данном направлении. В силу этого процессы столкновения называются также процессами рассеяния.

Пусть при столкновении электрон переходит из состояния (\mathbf{r}, \mathbf{k}) в состояние $(\mathbf{r}', \mathbf{k}')$. Поскольку при соударении координата не испытывает резкого изменения, то вероятность перехода в единицу времени не будет зависеть от \mathbf{r} и \mathbf{r}' , и пусть она равна $W(\mathbf{k}, \mathbf{k}')$. Это значит, что каждое столкновение сопровождается переходом электрона из состояния \mathbf{k} в состояние \mathbf{k}' . Однако, чтобы такой переход состоялся, состояние \mathbf{k}' должно быть свободным. 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 объеме $d\Gamma$, определяется вероятностью перехода $W(\mathbf{k}, \mathbf{k}')$, числом занятых исходных состояний $f(\mathbf{r}, \mathbf{k}, t) d\tau_{\mathbf{k}}/4\pi^3$ и числом свободных конечных состояний $[1 - f(\mathbf{r}, \mathbf{k}', t)] d\tau_{\mathbf{k}'}/4\pi^3$. Оно будет равно:

$$-f(\mathbf{r}, \mathbf{k}, t) W(\mathbf{k}, \mathbf{k}') [1 - f(\mathbf{r}, \mathbf{k}', t)] \frac{d\tau_{\mathbf{k}'}}{4\pi^3} \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (4-9)$$

В то же время существует конечная вероятность обратных переходов электронов из состояния \mathbf{k}' в состояние \mathbf{k} . Число электронов, совершивших обратный переход, сопровождающийся увеличением количества электронов в объеме $d\Gamma$, определяется равенством

$$f(\mathbf{r}, \mathbf{k}', t) W(\mathbf{k}', \mathbf{k}) [1 - f(\mathbf{r}, \mathbf{k}, t)] \frac{d\tau_{\mathbf{k}}}{4\pi^3} \frac{d\tau_{\mathbf{k}'}}{4\pi^3} d\tau_{\mathbf{r}} dt, \quad (4-10)$$

где $W(\mathbf{k}', \mathbf{k})$ — вероятность перейти электрону за 1 сек из состояния \mathbf{k}' в состояние \mathbf{k} ; член $[1 - f(\mathbf{r}, \mathbf{k}, t)] \frac{d\tau_{\mathbf{k}}}{4\pi^3}$ определяет наличие

свободного к состоянию. Следовательно, в результате соударений количество электронов в элементе объема фазового пространства за время dt изменится на величину

$$\{f(\mathbf{k}') W(\mathbf{k}', \mathbf{k}) [1 - f(\mathbf{k})] - f(\mathbf{k}) W(\mathbf{k}, \mathbf{k}') [1 - f(\mathbf{k}')]\} \frac{d\tau_{\mathbf{k}'}}{4\pi^3} \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (4-11)$$

Здесь ради сокращения записи f рассматривается как функция волнового вектора.

Для того чтобы найти полное число электронов, рассеянных в результате соударений, необходимо учесть всевозможные изменения \mathbf{k}' , т. е. необходимо выражение (4-11) проинтегрировать по объему зоны Бриллюэна

$$\frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt \int_{(V_B)} \{f(\mathbf{k}') W(\mathbf{k}', \mathbf{k}) [1 - f(\mathbf{k})] - f(\mathbf{k}) W(\mathbf{k}, \mathbf{k}') [1 - f(\mathbf{k}')]\} \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (4-12)$$

Таким образом, движение электронов и действие на них внешней силы, а также наличие процессов рассеяния приводит к изменению числа электронов в элементе объема фазового пространства. Это изменение за время от t до $t + dt$ составляет величину, равную

$$f(\mathbf{r}, \mathbf{k}, t + dt) \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} - f(\mathbf{r}, \mathbf{k}, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} = \frac{\partial f}{\partial t} \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (4-13)$$

Следовательно, выражение (4-13) равно сумме выражений (4-6), (4-7) и (4-12). После сокращения на $\frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt$ получим:

$$\frac{\partial f}{\partial t} = -(\mathbf{v} \nabla_{\mathbf{r}} f) - \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) + \int_{(V_B)} \{f(\mathbf{k}') W(\mathbf{k}', \mathbf{k}) [1 - f(\mathbf{k})] - f(\mathbf{k}) W(\mathbf{k}, \mathbf{k}') [1 - f(\mathbf{k}')]\} \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (4-14)$$

Согласно принципу микроскопической обратимости вероятности прямых и обратных переходов равны, т. е.

$$W(\mathbf{k}, \mathbf{k}') = W(\mathbf{k}', \mathbf{k}). \quad (4-15)$$

С учетом (4-15) соотношение (4-14) упрощается:

$$\frac{\partial f}{\partial t} = -(\mathbf{v} \nabla_{\mathbf{r}} f) - \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) + \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (4-16)$$

Уравнение (4-16) называется кинетическим уравнением Больцмана. Это уравнение является интегро-дифференциальным, так как под интегралом стоит искомая функция.

Легко видеть физический смысл уравнения (4-16). Левая часть его отражает изменение функции распределения во времени. Первый и второй члены правой части уравнения показывают изменения функции распределения, создаваемые внешними полями, приводящими к движению электронов в обычном пространстве и в пространстве волнового вектора. Изменение функции распределения электронов по состояниям в результате их рассеяния (столкновения) учтено в уравнении последним членом. Поэтому уравнение (4-16) можно записать в виде

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{\text{пол}} + \left(\frac{\partial f}{\partial t}\right)_{\text{ст}}. \quad (4-17)$$

Здесь член

$$\left(\frac{\partial f}{\partial t}\right)_{\text{пол}} = -(\mathbf{v}\nabla_{\mathbf{r}}f) - \frac{1}{\hbar}(\mathbf{F}\nabla_{\mathbf{k}}f) \quad (4-18)$$

носит название полевого члена уравнения Больцмана, а величина

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] d\tau_{\mathbf{k}'} \quad (4-19)$$

называется интегралом столкновений (или соударений).

В случае стационарных процессов, когда величины, описывающие явления переноса, не зависят от времени, функция распределения удовлетворяет равенству

$$\frac{\partial f}{\partial t} = 0. \quad (4-20)$$

Следовательно, кинетическое уравнение Больцмана для стационарного состояния имеет вид:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{пол}} = -\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} \quad (4-21)$$

или

$$(\mathbf{v}\nabla_{\mathbf{r}}f) + \frac{1}{\hbar}(\mathbf{F}\nabla_{\mathbf{k}}f) = \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (4-22)$$

Из выражений (4-21) и (4-22) следует, что в стационарном состоянии изменения функции распределения, вызванные движением носителей заряда и действием на них внешней силы, компенсируются соударениями носителей 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 их длины волны. Кроме того, его можно использовать для описания явления переноса в магнитном поле в том случае, если это поле не меняет спектра носителей в кристалле, т. е. не становятся существенными квантовые эффекты. Это условие накладывает ограничение на величину поля.

4-2. РАВНОВЕСНОЕ СОСТОЯНИЕ

При термодинамическом равновесии, при котором состояние электрона описывается равновесной функцией распределения (3-27), движение электрона в кристалле обуславливается тепловой энергией. Поэтому для равновесного случая левая часть уравнения (4-22) равна нулю. Это значит, что

$$\int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f_0(\mathbf{k}') - f_0(\mathbf{k})] d\tau_{\mathbf{k}'} = 0. \quad (4-23)$$

Из этого соотношения следует, что в системе, находящейся в равновесном состоянии, имеет место равенство потоков электронов для прямого и обратного процессов, что согласуется с принципом детального равновесия.

На основании выражения (4-23) получаем:

$$f_0(\mathbf{k}') = f_0(\mathbf{k}). \quad (4-24)$$

Подставляя в уравнение (4-24) выражение для равновесной функции распределения (3-27), будем иметь:

$$E_1 - F_1 = E_2 - F_2, \quad (4-25)$$

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 $E_1 = E_2$, откуда следует, что

$$F_1 = F_2. \quad (4-26)$$

Таким образом, в равновесном состоянии во всех частях системы, между которыми может происходить переход электронов, положение уровня Ферми одинаково.

4-3. ВРЕМЯ РЕЛАКСАЦИИ

Решение кинетического уравнения (4-16) в общем виде или даже выражения (4-21) для стационарного случая является весьма сложной задачей. Но она значительно упрощается для тех случаев, когда можно вести время релаксации. Как оказывается, описание процессов рассеяния при помощи времени релаксации возможно, если при столкновении энергия носителя заряда мало изменяется и если процессы рассеяния приводят к случайному распределению скоростей.

Пусть 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 процесса соударений система придет в равновесное состояние, описываемое равновесной функцией распределения f_0 . Это значит, что после выключения поля функция распределения изменяется вследствие соударений электронов с локальными нарушениями решетки

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{ст}. \quad (4-27)$$

В том случае, когда отклонение распределения носителей заряда от равновесного состояния невелико, можно положить, что в отсутствие внешних полей скорость изменения функции распределения f вследствие соударений пропорциональна величине ее отклонения от равновесной, т. е. пропорциональна $f - f_0$:

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{ст} = -\frac{f - f_0}{\tau(\kappa)}, \quad (4-28)$$

где $1/\tau(\kappa)$ — коэффициент пропорциональности, зависящий от κ . Решение уравнения (4-28) имеет вид:

$$f - f_0 = (f - f_0)_{t=0} e^{-t/\tau}. \quad (4-29)$$

Из (4-29) 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, величина τ есть среднее время, в течение которого в системе существует неравновесное распределение носителей заряда после выключения внешних полей. Поскольку установление равновесного состояния происходит в результате столкновений и при этом достаточно нескольких соударений, то время релаксации должно быть порядка времени свободного пробега электрона. С учетом этого можно считать, что длина свободного пробега l определяется скоростью движения электрона v и временем релаксации $\tau(\kappa)$.

Используя уравнение (4-28), интеграл столкновения (4-19) можно представить в виде

$$\left(\frac{\partial f}{\partial t} \right)_{ст} = \int_{(v_B)} W(\kappa, \kappa') [f(\kappa') - f(\kappa)] \frac{d\tau_{\kappa'}}{4\pi^3} = -\frac{f(\kappa) - f_0(\kappa)}{\tau(\kappa)}. \quad (4-30)$$

Поскольку внешние силы не приводят к большим изменениям равновесной функции, что определяется равенством (4-28), то неравновесную функцию распределения можно представить как

$$f(\kappa) = f_0 + f_1(\kappa), \quad (4-31)$$

где $f_1(\kappa)$ — поправка к равновесной функции распределения, появившаяся в результате внешнего воздействия на систему. Запишем

ее в виде

$$f_1(\kappa) = - \frac{\partial f_0}{\partial E} \kappa \chi(E). \quad (4-32)$$

Будем считать, что время релаксации не зависит от внешних полей. Это значит, что $\tau(\kappa)$ определяет и процессы рассеяния при наличии внешних полей, поэтому кинетическое уравнение (4-22) для стационарного состояния на основании формулы (4-31) можно записать в виде

$$(\mathbf{v} \nabla_{\mathbf{r}} f) + \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) = - \frac{f - f_0}{\tau(\kappa)} = - \frac{f_1(\kappa)}{\tau(\kappa)}. \quad (4-33)$$

Интеграл столкновения (4-19) с учетом формулы (4-31) принимает вид:

$$\left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = \frac{1}{4\pi^3} \int_{(V_B)} W(\kappa, \kappa') [f_1(\kappa') - f_1(\kappa)] d\tau_{\kappa'}. \quad (4-34)$$

Из выражения (4-30) можно определить время релаксации. Если учесть соотношения (4-34), (4-31) и (4-32), то $\tau(\kappa)$ будет определяться равенством следующего вида:

$$\begin{aligned} \frac{1}{\tau(\kappa)} &= - \frac{1}{4\pi^3} \int_{(V_B)} W(\kappa, \kappa') \frac{f_1(\kappa') - f_1(\kappa)}{f_1(\kappa)} d\tau_{\kappa'} = \\ &= \frac{1}{4\pi^3} \int_{(V_B)} W(\kappa, \kappa') \left[1 - \frac{\partial f_0 / \partial E' \kappa' \chi(E')}{\partial f_0 / \partial E \kappa \chi(E)} \right] d\tau_{\kappa'}. \end{aligned} \quad (4-35)$$

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 она после удара будет больше или меньше, чем до удара.

Будем считать, что соударения электронов со структурными нарушениями являются упругими. Это значит, что при упругом соударении изменяется лишь направление волнового вектора, и, следовательно, изменяется направление скорости движения электрона, а модули их остаются неизменными, т. е.

$$\kappa' = \kappa, \quad v' = v, \quad \text{а значит, } E' = E, \quad (4-36)$$

если зоны энергии сферические.

В силу этого время релаксации для упругого соударения с учетом равенства (4-36) будет равно:

$$\frac{1}{\tau(\kappa)} = \frac{1}{4\pi^3} \int_{(V_B)} W(\kappa, \kappa') \left[1 - \frac{\kappa' \chi(E')}{\kappa \chi(E)} \right] d\tau_{\kappa'}. \quad (4-37)$$

Пусть векторы χ , κ' и κ направлены так, как это изображено на рис. 4-2. Тогда проекции κ' и κ на вектор χ будут соответственно

равны:

$$\kappa_{\chi} = \kappa; \kappa'_{\chi} = \kappa' \cos \theta = \kappa \cos \theta, \quad (4-38)$$

где θ — угол отклонения электрона от первоначального направления движения в результате столкновения. Тогда равенство (4-37) с учетом равенства (4-38) запишется в виде

$$\begin{aligned} \frac{1}{\tau(\kappa)} &= \frac{1}{4\pi^3} \int_{(V_B)} W(\kappa, \kappa') \left[1 - \frac{\kappa'_{\chi}}{\kappa_{\chi}} \right] d\tau_{\kappa'} = \\ &= \frac{1}{4\pi^3} \int_{(V_B)} W(\kappa, \kappa') (1 - \cos \theta) d\tau_{\kappa'}. \end{aligned} \quad (4-39)$$

Таким образом, при рассмотрении процессов переноса основной задачей является вычисление неравновесной функции распределения, с помощью которой можно рассчитать все величины, характеризующие кинетические эффекты, например, плотность электрического тока, тепловой поток, напряженность электрического поля. Но, как следует из равенства (4-33), эта задача решается при условии, если известно время релаксации (4-39), а следовательно, необходимо знание механизма рассеяния носителей заряда.

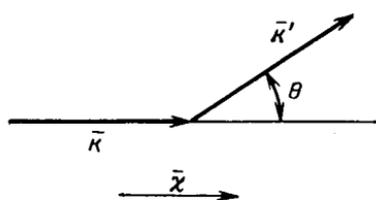


Рис. 4-2. Изменение волнового вектора электрона при рассеянии.

Рассмотрим теперь рассеяние, приводящее к случайному распределению скоростей (т. е. к такому, при котором носители заряда рассеиваются под углом 180°). В этом случае электрон, рассеиваясь, попадает в состояние κ' , которое находится возле того же минимума энергии, что и κ , но с противоположной скоростью. Но поскольку $W(\kappa, \kappa') = W(\kappa', \kappa)$ — четная функция, а $f_1(\kappa) = -\frac{\partial f_0}{\partial E} \kappa'_{\chi}(E)$ — нечетная функция относительно инверсии $\kappa \rightleftharpoons -\kappa$, поэтому

$$\int_{(V_B)} W(\kappa, \kappa') f_1(\kappa') d\tau_{\kappa'} = - \int_{(V_B)} \frac{\partial f_0}{\partial E} W(\kappa, \kappa') (\kappa'_{\chi}) d\tau_{\kappa'} = 0, \quad (4-40)$$

так как нечетная функция при интегрировании в симметричных пределах дает нуль. С учетом уравнения (4-40) интеграл столкновения (4-34) будет равен:

$$\left(\frac{\partial f}{\partial t} \right)_{ст} = - \frac{1}{4\pi^3} \int_{(V_B)} W(\kappa, \kappa') f_1(\kappa) d\tau_{\kappa'} = - \frac{f_1(\kappa)}{\tau(\kappa)}, \quad (4-41)$$

откуда следует, что время релаксации при рассеянии, приводящем к случайному распределению скоростей, определяется выражением:

$$\frac{1}{\tau(\kappa)} = \frac{1}{4\pi^3} \int_{(V_B)} W(\kappa, \kappa') d\tau_{\kappa'}. \quad (4-42)$$

4-4. ЭФФЕКТИВНОЕ СЕЧЕНИЕ РАССЕЯНИЯ. ТИПЫ ЦЕНТРОВ РАССЕЯНИЯ

Любая неоднородность кристалла, искажающая периодичность поля решетки, является центром рассеяния. Поскольку природа дефектов различна, то и рассеяние носителей заряда на них будет различно. Для количественной оценки процесса рассеяния вводится величина, называемая эффективным сечением.

Предположим, что имеется n свободных электронов, которые со средней скоростью v_0 движутся в данном направлении. Тогда nv_0 есть плотность потока электронов, т. е. число электронов, проходящих за единицу времени через единичное сечение. Пусть на пути этого потока на единичное сечение имеется N одинаковых центров рассеяния. Каждый центр характеризуется эффективным сечением, равным σ . Это по существу то пространство около центра, в области которого имеет место рассеяние электронов. Поэтому число рассеянных электронов n_1 в единицу времени определяется эффективным сечением рассеяния, количеством центров рассеяния и плотностью падающего потока электронов, т. е.

$$n_1 = \sigma N n v_0. \quad (4-43)$$

С другой стороны, если W — вероятность рассеяния одной частицы в единицу времени, то количество рассеянных электронов за 1 сек будет равно:

$$n_1 = W n. \quad (4-43a)$$

На основании равенств (4-43) и (4-43a) можно записать:

$$\sigma = \frac{n_1/N}{n v_0} = \frac{W}{N v_0}. \quad (4-44)$$

Таким образом, эффективное сечение рассеяния σ есть отношение числа удаленных электронов из пучка в результате рассеяния на одном центре в единицу времени к плотности падающего пучка.

Эффективное сечение рассеяния имеет размерность площади

$$[\sigma] = \frac{[W]}{[N v_0]} = \frac{T^{-1}}{L^{-3} L T^{-1}} = L^2.$$

Из формулы (4-44) найдем, что вероятность рассеяния равна:

$$W = \sigma N v_0. \quad (4-45)$$

Следовательно, вероятность рассеяния или вероятность столкновения определяется эффективным сечением центра, их количеством и скоростью движения носителя заряда.

В то же время вероятность столкновения обратно пропорциональна времени свободного пробега

$$W = \frac{1}{\tau}. \quad (4-46)$$

Поэтому время свободного пробега и длину свободного пробега можно выразить через эффективное сечение. Из формул (4-45) и

(4-46) следует, что

$$\tau = \frac{1}{\sigma N v_0} = \frac{l}{v_0} \quad (4-47)$$

и

$$l = \frac{1}{\sigma N}. \quad (4-48)$$

Величина $l^{-1} = \sigma N$ есть вероятность рассеяния на единичном интервале пути.

Рассмотрим случай, когда существует несколько процессов рассеяния. Пусть i -й механизм характеризуется эффективным сечением σ_i и числом рассеивающих центров N_i , и он определяет длину свободного пробега l_i . Если все эти процессы рассеяния возможны, т. е. может осуществиться любой из них, то согласно теории вероятности полная вероятность рассеяния в единицу времени будет определяться суммой отдельных вероятностей рассеяния:

$$W = \sum W_i. \quad (4-49)$$

Так как $\sum_i \sigma_i N_i = \sum_i l_i^{-1}$, то полная длина свободного пробега может быть определена из соотношения

$$\frac{1}{l} = \sum_i \frac{1}{l_i}. \quad (4-50)$$

Следовательно, полная длина свободного пробега всегда меньше самой малой парциальной длины свободного пробега.

В полупроводниках возможно несколько типов дефектов, нарушающих периодичность кристалла. Центрами рассеяния в них могут быть тепловые колебания решетки и статические дефекты, такие как атомы и ионы примеси, вакансии, междуузельные атомы, дислокации, дефекты упаковки, границы двойников и кристаллитов. Роль этих центров в рассеянии различна, а поэтому они должны иметь разное эффективное сечение. Для их количественной оценки за эффективное сечение рассеяния σ примем величину площадки, в пределах которой возможно взаимодействие между носителем заряда и дефектом.

Такие дефекты, как вакансии, междуузельные атомы, во многих отношениях сходны с примесями замещения. Эти дефекты называются точечными дефектами. Для них за величину 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 равна $N_A = 10^{16} \text{ см}^{-3}$, то $l_A = (3 \cdot 10^{-15} \cdot 10^{16})^{-1} \approx 3 \cdot 10^{-2} \text{ см}$.

Для иона примеси будем считать, что его диаметр в 10 раз больше диаметра примесного центра, т. е. $\sigma_I \approx (5 \cdot 10^{-8} \cdot 10)^2 \approx 3 \cdot 10^{13} \text{ см}^2$. 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, распространяющимися на большие области кристалла. Предположим, что линейная дислокация имеет длину 1 мм, а ее диаметр измеряется сотней периодов решетки, тогда площадь ее осевого сечения равна $5 \cdot 10^{-8} \cdot 100 \cdot 10^{-1} = 5 \cdot 10^{-7} \text{ см}^2$. При объемной плотности дислокаций $N_D = 10^8 \text{ см}$ длина свободного пробега будет порядка $l_D = 2 \cdot 10^{-2} \text{ см}$.

Границы двойников и кристаллитов, 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 $r = 1 \text{ \AA}$, а диаметр атома $d \approx 10^{-8} \text{ см}$, то $\sigma_T = (d + r)^2 - d^2 \approx 2dr = 10^{-16} \text{ см}^2$. Эта величина значительно меньше, чем для других центров рассеяния, но число колеблющихся атомов велико: $N_T \approx 10^{22} \text{ см}^{-3}$, поэтому $l_T \approx 10^{-8} \text{ см} = 100 \text{ \AA}$.

4.5. РАССЕЯНИЕ НА ИОНАХ ПРИМЕСИ

В примесном полупроводнике каждый ион примеси создает вокруг себя электрическое поле. Под воздействием этого поля носитель заряда отклоняется от своего первоначального направления (рис. 4-3). При этом носитель отклоняется тем сильнее, чем медленнее он движется и ближе проходит мимо иона примеси.

Задача сводится фактически к нахождению траектории движе-

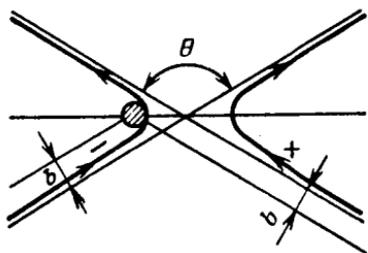


Рис. 4-3. Искривление траектории движения электрона в поле положительного иона примеси.

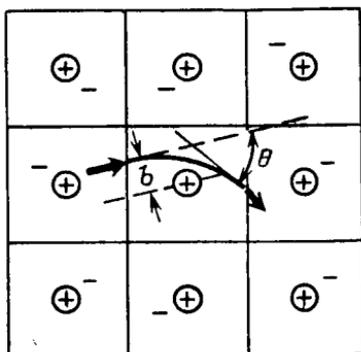


Рис. 4-4. Рассеяние электрона и дырки на положительном ионе примеси.

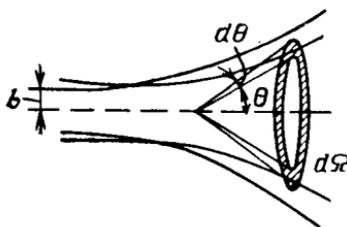
ния заряженной частицы в центральном поле иона. Она может быть решена чисто классическими методами. Для этого будем считать носитель заряда классической частицей, движущейся в кулоновском поле положительного иона примеси с зарядом Ze . В этом случае потенциальную энергию взаимодействия можно представить

в виде

$$U(r) = \pm \frac{Ze^2}{er} \quad (\text{в системе Гаусса}). \quad (4-51)$$

Плюс соответствует дырке, минус — электрону.

Как показывает расчет¹, в кулоновском поле иона носитель заряда движется по гиперболе. На рис. 4-4 изображены траектории электрона и дырки в поле положительного иона. Оба носителя заряда одинаково отклоняются, хотя в одном случае действует сила притяжения, в другом — отталкивания. Через b обозначено прицельное расстояние — минимальное расстояние между ионом и траекторией носителя заряда, а θ — угол рассеяния, т. е. угол между первоначальным направлением движения носителя и его движением после рассеяния. При этом угол рассеяния зависит от прицельного расстояния согласно соотношению вида:



$$b = \frac{Ze^2}{em^*v^2} \operatorname{ctg} \frac{\theta}{2}, \quad (4-52)$$

где m^* и v — эффективная масса и скорость движения носителя заряда.

Рис. 4-5. Телесный конический угол рассеяния.

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.

Пусть 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. 4-5. Здесь через $d\Omega$ обозначен телесный угол, заключенный между этими конусами. При этом он равен:

$$d\Omega = 2\pi \sin \theta d\theta. \quad (4-53)$$

Процессы рассеяния в данном случае характеризуются дифференциальным эффективным сечением $\sigma(\theta)$. По определению $\sigma(\theta)$ равно:

$$\sigma(\theta) = \frac{\text{число частиц, отклоненных одним центром в 1 сек}}{\text{на угол } \theta \text{ в единичный телесный угол}} \cdot \frac{\text{поток падающих частиц на 1 см}^2 \text{ в 1 сек}}{\text{на 1 см}^2 \text{ в 1 сек}}. \quad (4-54)$$

Пусть имеется n электронов, движущихся хаотически со скоростями v . Если выбрать перпендикулярно каждому направлению движения электронов площадку в 1 см^2 , то на нее за 1 сек упадет nv электронов. Полное число частиц, отклоненных за 1 сек в телес-

¹ Борн М., Современная физика, ОНТИ, Л.—М., 1935, стр. 196.

ный угол $d\Omega$, будет равно потоку частиц, падающих на кольцо, согласно рис. 4-5, площадью $2\pi b |db|$. Поэтому на основании определения (4-54) и равенства (4-53) можно записать:

$$\sigma(\theta) = \frac{nv2\pi b |b|}{nv d\Omega} = \frac{b |db|}{\sin \theta d\theta}. \quad (4-55)$$

Из формулы (4-52) имеем:

$$|db| = \frac{Ze^2}{2\epsilon m^* v^2} \frac{1}{\sin^2 \frac{\theta}{2}} d\theta. \quad (4-56)$$

Подставив формулы (4-52) и (4-56) в (4-55), получим:

$$\sigma(\theta) = \left(\frac{Ze^2}{2\epsilon m^* v^2} \right)^2 \frac{1}{\sin^4 \frac{\theta}{2}}. \quad (4-57)$$

Формула (4-57) была получена Резерфордом при исследовании рассеяния α -частиц полем ядер.

Для подсчета времени релаксации необходимо выразить вероятность перехода $W(\kappa, \kappa')$ через дифференциальное сечение рассеяния. Предположим, что в кристалле единичного объема имеется N_I ионов примеси, тогда число рассеянных носителей всеми центрами в 1 сек в телесный угол $d\Omega$ составит величину

$$nW(\kappa, \kappa') d\Omega = \sigma(\theta) d\Omega nvN_I, \quad (4-58)$$

откуда

$$W(\kappa, \kappa') = \sigma(\theta) vN_I. \quad (4-59)$$

Таким образом, вероятность перехода электрона из состояния κ в состояние κ' 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 электрона в 10^4 раз меньше массы иона, являющегося рассеивающим центром. Во-вторых, ионы решетки гораздо менее подвижны относительно своего положения равновесия, чем носители заряда. Поэтому при соударении носителя заряда с ионом примеси энергия носителя изменяется очень незначительно.

Если при упругом соударении носитель рассеивается в телесный угол $d\Omega$, то время релаксации, используя (4-39), можно записать в виде

$$\frac{1}{\tau(\kappa)} = \int_0^\pi W(\kappa, \kappa') [1 - \cos \theta] d\Omega. \quad (4-60)$$

Подставив формулу (4-59) в (4-60), получим:

$$\frac{1}{\tau(\kappa)} = 2\pi N_I v(\kappa) \int_0^\pi \sigma(\theta) (1 - \cos \theta) \sin \theta d\theta = N_I v(\kappa) \sigma_c. \quad (4-61)$$

Величина

$$\sigma_c = 2\pi \int_0^{\pi} \sigma(\theta) (1 - \cos \theta) \sin \theta d\theta \quad (4-62)$$

есть усредненное дифференциальное эффективное сечение; усреднение по углам θ производится с весовой функцией $(1 - \cos \theta)$. σ_c называется эффективным сечением проводимости или подвижности, а также транспортным эффективным сечением.

Время релаксации на основании формулы (4-61) равно:

$$\tau(\kappa) = \frac{1}{N_I v(\kappa) \sigma_c}, \quad (4-63)$$

а длина свободного пробега

$$l(\kappa) = \tau(\kappa) v(\kappa) = \frac{1}{N_I \sigma_c}. \quad (4-64)$$

Из формулы (4-57) 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 (4-62) подставить значение $\sigma(\theta)$ из формулы (4-57), то интеграл на нижнем пределе при $\theta = 0$ будет расходящимся и сечение рассеяния будет бесконечным. Рассеяние носителей заряда на ионизированной примеси обладает конечным сечением, поэтому при определении времени релаксации интегрирование выражения (4-61) проводится от $\theta_{\text{мин}}$ до π . Определить $b_{\text{макс}}$ 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 $b_{\text{макс}} = 1/2 N_I^{-1/3}$. Он соответствует наименьшему значению угла рассеяния $\theta_{\text{мин}}$, который можно определить из формулы (4-52):

$$\text{ctg} \frac{\theta_{\text{мин}}}{2} = \frac{em^*v^2}{Ze^2} b_{\text{макс}} = \frac{em^*v^2}{2Ze^2 N_I^{1/3}}. \quad (4-65)$$

С учетом (4-65) равенство (4-61) примет вид:

$$\begin{aligned} \frac{1}{\tau(\kappa)} &= 2\pi N_I v(\kappa) \int_{\theta_{\text{мин}}}^{\pi} \sigma(\theta) (1 - \cos \theta) \sin \theta d\theta = \\ &= 8\pi \left(\frac{Ze^2}{2em^*v^2} \right)^2 N_I v(\kappa) \ln \frac{1}{\sin^2 \frac{\theta_{\text{мин}}}{2}} = \\ &= 2\pi \left(\frac{Ze^2}{em^*v^2} \right)^2 N_I v(\kappa) \ln \left[1 + \text{ctg}^2 \frac{\theta_{\text{мин}}}{2} \right]. \end{aligned} \quad (4-66)$$

Таким образом, время релаксации при рассеянии носителей заряда на ионизированных атомах примеси равно:

$$\tau(\kappa) = \frac{e^2 m^{*2} v^3(\kappa)}{2\pi Z^2 e^4 N_I \ln \left[1 + \left(\frac{em^* v^2}{2Ze^2 N_I^{1/3}} \right)^2 \right]} \quad (4-67)$$

Это соотношение для τ называется формулой Конуэлл-Вайскопфа.

Если выразить скорость через энергию $v = \left(\frac{2E}{m^*} \right)^{1/2}$, то время релаксации носителей заряда при рассеянии на ионах примеси примет вид:

$$\tau_I = \tau_{0I} E^{3/2}, \quad (4-68)$$

где

$$\tau_{0I} = \frac{(2m^*)^{1/2} e^3}{\pi Z^2 e^4 N_I \ln \left[1 + \left(\frac{eE}{Ze^2 N_I^{1/3}} \right)^2 \right]}, \quad (4-69)$$

а длина свободного пробега равна:

$$l = v(\kappa) \tau(\kappa) = \left(\frac{2}{m^*} \right)^{1/2} \tau_{0I} E^2. \quad (4-70)$$

4-6. РАССЕЙНИЕ 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 ионизированных атомов примеси меньше концентрации нейтральных атомов примеси, этот механизм рассеяния играет существенную роль. В этом случае рассеяние носителей может осуществляться двумя путями: с помощью прямого упругого соударения и обменом падающего электрона с электроном примесного атома. Определение времени релаксации в этом случае сложно. Однако можно провести элементарный расчет, если рассеяние носителей заряда на нейтральных атомах примеси рассматривать как рассеяние медленных электронов на атоме водорода, находящегося в среде с диэлектрической проницаемостью ϵ . При этом была получена формула

$$\tau_A = \frac{em^{*2}}{20e\hbar^3} \frac{1}{N_A}, \quad (4-71)$$

где N_A — концентрация атомов примеси.

Как следует из формулы (4-71), 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 ионизирована.

Рассеяние носителей заряда может происходить на нарушениях решетки, вносимых дислокациями. Дислокацию в полупроводнике электронного типа можно уподобить линейному отрицательному заряду в виде бесконечно длинного цилиндра радиусом R , вокруг которого имеется положительный пространственный заряд. Электроны, взаимодействуя с отрицательно заряженной дислокацией, испытывают отталкивание, что и приводит к их рассеянию. 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

$$\tau_D = \frac{3}{8Rv} \frac{1}{N_D}. \quad (4-72)$$

Если принять $v = 10^7$ см/сек, $R = 3 \cdot 10^{-5}$ см, $N_D = 10^6$ см⁻², то время релаксации $\tau_D = 1,25 \cdot 10^{-9}$ сек. Эта величина на три порядка больше, чем время релаксации для рассеяния на тепловых колебаниях решетки при комнатной температуре, поэтому это рассеяние незначительно при комнатных температурах, но зато оно может быть существенно при низких температурах.

4-7. ТЕПЛОВЫЕ КОЛЕБАНИЯ АТОМОВ ОДНОМЕРНОЙ РЕШЕТКИ

Атомы кристаллической решетки вследствие теплового движения совершают хаотические колебания относительно своих положений равновесия. Это нарушает строгую периодичность поля решетки, что приводит к рассеянию на этих нерегулярностях носителей заряда.

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

Прежде чем рассматривать тепловое движение атомов одномерной решетки, проанализируем колебания одномерной однородной струны. Как известно, в упругой струне могут быть колебания двух типов: продольные, в которых движение каждого из элементов струны происходит в направлении ее длины, и поперечные, при которых направление движения части струны перпендикулярно направлению распространения колебаний.

Рассмотрим силы, действующие на элемент Δx однородной струны, имеющей линейную плотность ρ . Пусть для координаты x величина продольного смещения элемента из положения равновесия равна u . Если на одном конце этого элемента величина деформации e , определяемая как относительное изменение длины $e = \frac{du}{dx}$, то на другом конце элемента она будет $e(x + \Delta x) = e(x) + \frac{d^2u}{dx^2} \Delta x$. Если сила, вызывающая деформацию, равна F , то упругая постоян-

ная (модуль Юнга) будет $T_n = F/e$. В соответствии с этим результирующая сила, действующая на элемент Δx , будет $F_{\Delta x} = T_n \frac{d^2u}{dx^2} \Delta x$. С другой стороны, эту силу можно определить как произведение массы элемента $\rho \Delta x$ на ускорение $\frac{d^2u}{dt^2}$, т. е. $F_{\Delta x} = \rho \frac{d^2u}{dt^2} \Delta x$. В результате получим уравнение свободных колебаний струны

$$\frac{d^2u}{dt^2} = \frac{T_n}{\rho} \frac{d^2u}{dx^2}, \quad (4-73)$$

где скорость распространения упругой звуковой волны

$$v_{зв} = \sqrt{\frac{T_n}{\rho}} \quad (4-74)$$

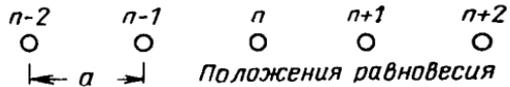
не зависит от частоты колебаний.

Решение уравнения колебаний струны имеет вид бегущей монокроматической волны

$$u = A \sin 2\pi \left(\frac{x}{\lambda} - \nu t \right) = A \sin (qx - \omega t) = Ae^{i(qx - \omega t)},$$

где A — амплитуда колебаний; ν — частота колебаний; ω — циклическая частота колебаний; λ — длина волны, $q = 2\pi/\lambda$ — волновое число.

Для упругой волны в непрерывной струне частота колебаний



$$\omega = 2\pi\nu = 2\pi \frac{v_{зв}}{\lambda} = v_{зв}q \quad (4-75)$$



пропорциональна волновому числу q . При этом абсолютная величина волнового числа q может принимать все значения от 0 до ∞ , и каждому значению q соответствует определенная форма волны. Частота колебаний также может изменяться от нуля до бесконечности.

Рис. 4-6. Схематическое изображение линейной решетки одинаковых атомов.

Рассмотрим теперь линейную цепочку одинаковых атомов массы m , расположенных на расстоянии a друг от друга, как показано на рис. 4-6. Пусть n -й атом смещен из положения равновесия на расстояние u_n . Если отклонения атомов от положения равновесия малы, то силы взаимодействия можно рассматривать как квазиупругие, т. е. пропорциональные отклонению. Поэтому при учете взаимодействия лишь ближайших соседей, поскольку на атом n действуют две силы, направленные в противоположные стороны,

сила F_n будет равна:

$$F_n = -\beta(u_n - u_{n+1}) - \beta(u_n - u_{n-1}) = \beta(u_{n+1} + u_{n-1} - 2u_n), \quad (4-76)$$

где β — коэффициент квазиупругой силы.

Уравнение движения для этого атома имеет вид:

$$m \frac{d^2 u_n}{dt^2} = \beta(u_{n+1} + u_{n-1} - 2u_n). \quad (4-77)$$

Решение этого уравнения будем искать в виде

$$u_n = A e^{i(qan - \omega t)}. \quad (4-78)$$

Подставив (4-78) в (4-77), получим:

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

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

$$\omega = \pm 2(\beta/m)^{1/2} \sin \frac{aq}{2} = \pm \omega_m \sin \frac{aq}{2}, \quad (4-79)$$

где

$$\omega_m = 2(\beta/m)^{1/2}. \quad (4-80)$$

Таким образом, решения типа упругой волны (4-78) удовлетворяют уравнениям (4-77) для любого n , если частота колебаний волны связана с волновым числом q дисперсионным соотношением вида (4-79). Скорость распространения упругой волны вдоль линейной решетки

$$v = \frac{\omega \lambda}{2\pi} = \frac{\lambda}{\pi} (\beta/m)^{1/2} \sin \frac{\pi a}{\lambda} \quad (4-81)$$

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

Установим пределы изменения волнового числа q . Для этого рассмотрим, как изменится уравнение колебаний для новой волны с волновым числом $q' = q + \frac{2\pi}{a}g$, где g — целое положительное или отрицательное число. В этом случае

$$(u_n)_g = A e^{i(qan - \omega t)} e^{i2\pi n g} = A e^{i(qan - \omega t)} = u_n; \quad (4-82)$$

так как n — порядковый номер атома, то ng — целое число и $e^{i2\pi n g} = 1$. Из соотношения (4-82) 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

$$-\frac{\pi}{a} \leq q \leq \frac{\pi}{a}. \quad (4-83)$$

Отсюда можно сделать вывод, что область, в которой лежат волновые векторы q , совпадает с зоной Бриллюэна для волнового вектора электрона k . Поэтому значения q можно рассматривать в том же пространстве, что и k .

Для определения величины q поступим следующим образом. Пусть число атомов N в одномерной цепочке велико. Расположим их по окружности так, чтобы N -й и 1-й атом находились в равновесии на расстоянии постоянной решетки a . Так как число атомов N велико, то радиус окружности будет большой. Поскольку силы взаимодействия между атомами распространяются на расстояния

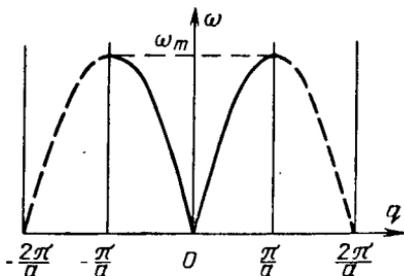


Рис. 4-7. Зависимость частоты колебаний от волнового числа для линейной решетки, состоящей из одинаковых атомов.

в пределах 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

$$u_{n \pm N} = u_n. \quad (4-84)$$

Если подставим равенство (4-78) в выражение (4-84), то получим

$$u_{n \pm N} = A e^{i[qa(n \pm N) - \omega t]} = A e^{i(qan - \omega t)} e^{\pm iqaN} = u_n. \quad (4-85)$$

Из равенства (4-85) следует, что $e^{\pm iqaN} = 1$. Это значит, что

$$qaN = 2\pi g, \quad (4-86)$$

откуда

$$q = \frac{2\pi}{aN} g. \quad (4-86a)$$

Учитывая неравенство (4-83), находим, что g должно меняться в пределах

$$-\frac{N}{2} \leq g \leq \frac{N}{2}. \quad (4-87)$$

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

Зависимость частоты ω от волнового числа q носит название закона дисперсии для колебаний атомов. Для линейной решетки одинаковых атомов закон дисперсии (положительная ветвь кривой) представлен на рис. 4-7.

Для больших значений длин волн, или малых значений q , можно положить $\sin \frac{aq}{2} \approx \frac{aq}{2}$. Тогда из (4-79) будем иметь:

$$\omega = \left(\frac{4\beta}{m}\right)^{1/2} \frac{aq}{2} = \sqrt{\frac{\beta}{m}} aq. \quad (4-88)$$

Но для линейной атомной цепочки линейная плотность $\rho = \frac{m}{a}$, а модуль Юнга

$$T_n = \frac{\text{сила}}{\text{относительное удлинение}} = \frac{F_n}{\frac{|u_n - u_{n-1}|}{a}} = \beta a,$$

откуда

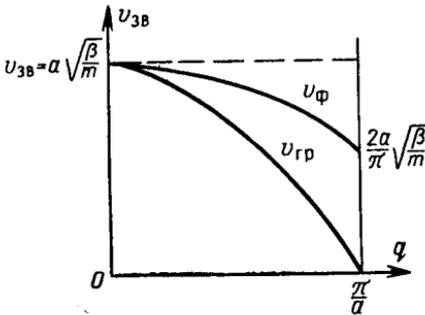
$$v_{зв} = \sqrt{\frac{T_n}{\rho}} = \sqrt{\frac{\beta}{m}} a. \quad (4-89)$$

Поэтому соотношение (4-88) с учетом равенства (4-89) примет вид:

$$\omega = v_{зв} q, \quad (4-90)$$

т. е. для длинных волн частота колебаний атомной цепочки пропорциональна волновому числу q .

Минимальная длина волны соответствует максимальному значению волнового числа $q_m = \pi/a$ и $\lambda_{\text{мин}} = 2a$. Физически это положение представляется ясным, так как в дискретной атомной цепочке не могут существовать волны с длиной волны меньше, чем период решетки a . Волне с наименьшей длиной $\lambda_{\text{мин}} = 2a$ соответствует максимальная частота, равная согласно (4-80) и (4-90)



$$\omega_m = 2 \left(\frac{\beta}{m}\right)^{1/2} = \frac{2v_{зв}}{a}. \quad (4-91)$$

Рис. 4-8. Зависимость фазовой и групповой скорости от волнового числа.

Таким образом, для колебаний дискретных атомных структур характерным является существование максимальной частоты. В твердых телах скорость звука порядка $3 \cdot 10^5$ см/сек, а постоянная решетки примерно $a \approx 3 \cdot 10^{-8}$ см, поэтому предельная частота равна $\omega_m = \frac{2 \cdot 3 \cdot 10^5}{3 \cdot 10^{-8}} \approx 2 \cdot 10^{13}$ гц и лежит в инфракрасной области.

Если имеет место дисперсия волн, то надо различать фазовую скорость, с которой распространяется фаза монохроматической волны, равную $v_\phi = \omega/q$, и групповую скорость, с которой распространяется волновой пакет, а следовательно, энергия волны

$v_{гр} = \frac{d\omega}{dq}$ (рис. 4-8). Для линейной цепочки одинаковых атомов согласно равенствам (4-79) и (4-89)

$$v_{ф} = \omega_m \left| \frac{\sin \frac{aq}{2}}{q} \right| = v_{зв} \left| \frac{\sin \frac{aq}{2}}{\frac{aq}{2}} \right|; \quad (4-92)$$

$$v_{гр} = \frac{a\omega_m}{2} \left| \cos \frac{aq}{2} \right| = v_{зв} \left| \cos \frac{aq}{2} \right|. \quad (4-93)$$

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

$$v_{ф} = v_{гр} = v_{зв}, \quad (4-94)$$

т. е. групповая и фазовая скорости равны и представляют собой скорость звука.

Для наиболее коротких длин волн, т. е. для $q_m = \pi/a$, $v_{ф} = \frac{2v_{зв}}{\pi} = \frac{2a}{\pi} \sqrt{\frac{\beta}{m}}$, а групповая скорость, с которой переносится энергия колебаний решетки, уменьшается до нуля.

В случае малых колебаний движение атома решетки можно описать с помощью линейной суперпозиции бегущих волн смещения типа (4-78). Каждая из этих волн будет характеризоваться волновым числом q , соответствующей ему частотой ω_q и амплитудой A_q . Отклонение n -го атома в этом случае можно представить в виде

$$u_n = \sum_q A_q e^{i(qan - \omega_q t)}. \quad (4-95)$$

Введем новые координаты, положив

$$Q_q = C_q e^{-i\omega_q t}, \quad (4-96)$$

где $C_q = \sqrt{N} A_q$.

Независимые переменные Q_q , изменяющиеся во времени по гармоническому закону, называются нормальными (или главными) координатами решетки кристалла. Подставив значения Q_q и A_q в (4-95), в результате будем иметь:

$$u_n = \frac{1}{\sqrt{N}} \sum_q Q_q e^{iqan}. \quad (4-97)$$

Определим теперь полную энергию колебаний линейной решетки. Она складывается из кинетической и потенциальной энергии всех атомов. Кинетическая энергия согласно (4-97) будет равна:

$$T = \sum_{n=1}^N \frac{m}{2} \dot{u}_n^2 = \frac{m}{2N} \sum_q \dot{Q}_q^2 \sum_{n=1}^N e^{i2qan}. \quad (4-98)$$

Подсчитаем сомножитель $\sum_{n=1}^N e^{i2qan}$, учитывая равенство (4-86а).

Он вычисляется как сумма геометрической прогрессии

$$\begin{aligned} \sum_{n=1}^N e^{i2qan} &= \sum_{n=1}^N e^{i\frac{4\pi}{N}gn} = e^{i\frac{4\pi}{N}g^1} + e^{i\frac{4\pi}{N}g^2} + \dots + e^{i\frac{4\pi}{N}g^N} = \\ &= \frac{e^{i\frac{4\pi}{N}g} - e^{i\frac{4\pi}{N}g} e^{i4\pi g}}{1 - e^{i\frac{4\pi}{N}g}} = \frac{e^{i\frac{4\pi}{N}g} (1 - e^{i4\pi g})}{1 - e^{i\frac{4\pi}{N}g}} = \\ &= \begin{cases} 0, & \text{когда } q \neq 0 \text{ (или } g \neq 0), \\ N, & \text{когда } q = 0 \text{ (или } g = 0), \end{cases} \end{aligned} \quad (4-99)$$

так как $e^{i4\pi g} = 1$, если целое число $g \neq 0$, а сумма ряда равна N , если $g = 0$, поскольку $\sum_{n=1}^N 1 = N$.

С учетом системы (4-99) выражение (4-98) запишется в виде

$$T = \sum_q \frac{m}{2} \dot{Q}_q^2. \quad (4-100)$$

По определению сила $F_n = -\frac{\partial U}{\partial u_n}$, где U — потенциальная энергия, поэтому потенциальная энергия решетки с учетом (4-76) и (4-78) будет равна:

$$U = U_0 + \frac{\beta}{2} \sum_{n=1}^N (2 - e^{iqa} - e^{-iqa}) u_n^2. \quad (4-101)$$

Для определения β используем уравнение движения (4-77), в которое подставим величины смещения (4-95):

$$m \sum_q \omega_q^2 A_q e^{i(qan - \omega_q t)} = \beta \sum_q (2 - e^{-iqa} - e^{iqa}) A_q e^{i(qan - \omega_q t)}, \quad (4-102)$$

откуда найдем

$$\beta = \sum_q \frac{m\omega_q^2}{2 - e^{iqa} - e^{-iqa}}. \quad (4-103)$$

Теперь в уравнение (4-101) подставим значение β и величины смещения (4-97), выраженные через нормальные координаты. В результате, используя систему (4-99), получаем:

$$U = U_0 + \sum_q \frac{m}{2} \omega_q^2 Q_q^2. \quad (4-104)$$

Если $U_0 = 0$, то

$$U = \sum_q \frac{m}{2} \omega_q^2 Q_q^2. \quad (4-104а)$$

Полная энергия колебаний решетки имеет вид:

$$E = T + U = \sum_q \left(\frac{m}{2} \dot{Q}_q^2 + \frac{m}{2} \omega_q^2 Q_q^2 \right). \quad (4-105)$$

Выражение

$$E_q = \frac{m}{2} \dot{Q}_q^2 + \frac{m}{2} \omega_q^2 Q_q^2 \quad (4-106)$$

совпадает с выражением полной энергии гармонического осциллятора с массой m . Спектр разрешенных значений энергии q -го квантовомеханического осциллятора записывается в виде

$$E_q = \hbar \omega_q \left(n_q + \frac{1}{2} \right) \quad (n_q = 0, 1, 2, \dots). \quad (4-107)$$

Из (4-105) следует, что полная энергия колебаний решетки, выраженная через нормальные координаты, имеет вид энергии совокупности гармонических осцилляторов. Это означает,

что кристалл, содержащий N атомов, можно рассматривать как динамическую систему с N степенями свободы и ее движение можно описать с помощью N нормальных координат. Каждая нормальная координата описывает определенную конфигурацию всех атомов кристалла и удовлетворяет уравнению

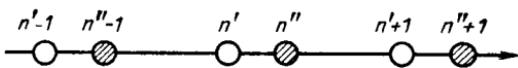


Рис. 4-9. Одномерная модель решетки, имеющей два атома в элементарной ячейке.

Каждая нормальная координата описывает определенную конфигурацию всех атомов кристалла и удовлетворяет уравнению

$$\ddot{Q}_q + \omega_q^2 Q_q = 0. \quad (4-108)$$

Такое колебание носит название **нормального колебания**.

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 . Поэтому энергия решетки может быть представлена в виде суммы энергии квантовомеханических гармонических осцилляторов:

$$E = \sum_q E_q = \sum_q \hbar \omega_q \left(n_q + \frac{1}{2} \right) \quad (n_q = 0, 1, 2, \dots). \quad (4-109)$$

Рассмотрим теперь одномерный кристалл, состоящий из атомов двух сортов с массами m_1 и m_2 (рис. 4-9). В качестве примеров веществ, у которых в элементарной ячейке имеется по два атома, можно привести ионные щелочно-галогидные соединения, ковалентные полупроводники — германий, кремний, а также полупроводниковые соединения типа $A^{III}B^V$.

Пусть атомы массы m_1 находятся в узлах решетки n' , а атомы массы m_2 — в узлах n'' . Обозначим смещение n' -го и n'' -го атомов через u_n и u_n'' и будем считать, что коэффициент квазиупругой силы

между соседними атомами одинаков и равен β . Уравнение движения атомов обоих сортов имеет вид:

$$m_1 \ddot{u}_n = \beta (u_n'' + u_{n-1}'' - 2u_n''); \quad (4-110)$$

$$m_2 \ddot{u}_n'' = \beta (u_n' + u_{n+1}' - 2u_n'').$$

Решение этих уравнений будем искать в виде

$$u_n^a = A_a e^{i(qan - \omega t)}. \quad (4-111)$$

Подставляя выражения (4-111) в (4-110), получаем:

$$\begin{aligned} (2\beta - m_1 \omega^2) A_1 &= \beta (1 + e^{-iqa}) A_2; \\ (2\beta - m_2 \omega^2) A_2 &= \beta (1 + e^{iqa}) A_1. \end{aligned} \quad (4-112)$$

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

$$\begin{vmatrix} 2\beta - m_1 \omega^2 & \beta (1 + e^{-iqa}) \\ \beta (1 + e^{iqa}) & 2\beta - m_2 \omega^2 \end{vmatrix} = 0. \quad (4-113)$$

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

$$\omega^4 - 2\beta \frac{m_1 + m_2}{m_1 m_2} \omega^2 - \frac{4\beta^2}{m_1 m_2} \sin^2 \frac{aq}{2} = 0. \quad (4-114)$$

Из этого уравнения найдем:

$$\omega_1^2 = \frac{\omega_0^2}{2} \left[1 + \sqrt{1 - \gamma^2 \sin^2 \frac{aq}{2}} \right] \quad (4-115)$$

и

$$\omega_2^2 = \frac{\omega_0^2}{2} \left[1 - \sqrt{1 - \gamma^2 \sin^2 \frac{aq}{2}} \right], \quad (4-116)$$

где $\omega_0^2 = 2\beta \frac{m_1 + m_2}{m_1 m_2}$ и $\gamma^2 = 4 \frac{m_1 m_2}{(m_1 + m_2)^2}$.

Из формул (4-115) и (4-116) 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 находятся два атома, имеются две ветви колебаний; одна из них называется оптической с частотой ω_1 , другая — акустической с частотой ω_2 .*

Рассмотрим два предельных случая: длинные и короткие волны. Длинные волны имеют место вблизи $q = 0$. В случае $q = 0$ частоты колебаний на основании (4-115) и (4-116) равны:

$$\omega_{оп}(0) = \left(2\beta \frac{m_1 + m_2}{m_1 m_2} \right)^{1/2} = \omega_0 \quad (4-117)$$

и

$$\omega_{ак}(0) = 0. \quad (4-118)$$

Для акустической волны колебаний согласно (4-111) и (4-112)

$$\left(\frac{u'_n}{u''_n}\right)_{\text{ак}} = \left(\frac{A_1}{A_2}\right)_{\text{ак}} = 1, \quad (4-119)$$

т. е. смещения обоих атомов одинаковы. 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. 4-10).

Аналогичное положение имеет место при распространении упругих волн. Поэтому эта ветвь колебаний и была названа акустической. Таким образом, при малых значениях q акустическая ветвь колебаний соответствует обычным звуковым волнам и скорость их распространения равна скорости звука в кристалле, а

$$\omega_{\text{ак}} = v_{\text{зв}} q. \quad (4-120)$$

Максимально длинной волне оптической ветви колебания соответствует частота, определяемая соотношением (4-117). Для нее с учетом (4-111) и (4-112) справедливо:

$$\left(\frac{u'_n}{u''_n}\right)_{\text{оп}} = \left(\frac{A_1}{A_2}\right)_{\text{оп}} = -\frac{m_2}{m_1}, \quad (4-121)$$

т. е. 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. 4-10). Если бы элементарную ячейку составляли ионы разного знака, то их колебания в противофазе привели бы к изменению электрического дипольного момента и, следовательно, к поглощению и испусканию инфракрасного излучения. Поэтому колебания этого типа и были названы оптическими.

Для длинных волн, а следовательно, для малых значений q , полагая $\sin \frac{aq}{2} \approx \frac{aq}{2}$ и разлагая корни в уравнениях (4-115) и (4-116) в ряд, имеем для оптических колебаний:

$$\omega_{\text{оп}} \approx \omega_0 \left(1 - \frac{\gamma^2 a^2}{32} q^2\right). \quad (4-122)$$

С ростом q частота оптических колебаний $\omega_{\text{оп}}$ уменьшается по квадратичному закону. При $q \rightarrow 0$ фазовая скорость оптических



Рис. 4-10. Схематическое изображение колебаний частиц в двухатомной линейной решетке, иллюстрирующее природу различия между оптическими и акустическими колебаниями. Предполагается, что колебания поперечные и что длина волны обоих типов колебаний одинакова.

колебаний стремится к бесконечности, а групповая скорость равна нулю.

Для длинных волн акустических колебаний

$$\omega_{\text{ак}} \approx \frac{1}{4} \omega_0 \gamma a q = \sqrt{\frac{2\beta}{m_1 + m_2}} a q = v_{\text{зв}} q, \quad (4-123)$$

а фазовая и групповая скорости равны скорости звука

$$v_{\text{ф}} = v_{\text{гр}} = v_{\text{зв}}. \quad (4-124)$$

Для самой короткой длины волны $\lambda_{\text{мин}} = 2a$, когда $q_m = \frac{\pi}{a}$, частота оптических колебаний будет минимальная

$$\omega_{\text{оп}} \left(\frac{\pi}{a} \right) = \left(\frac{2\beta}{m_2} \right)^{1/2} = \omega_{\text{оп}}^{\text{мин}}, \quad (4-125)$$

а частота акустических колебаний будет максимальная

$$\omega_{\text{ак}} \left(\frac{\pi}{a} \right) = \left(\frac{2\beta}{m_1} \right)^{1/2} = \omega_{\text{ак}}^{\text{макс}} \quad (4-126)$$

Зависимости ω от q для оптических и акустических ветвей колебаний одномерной решетки, имеющей в элементарной ячейке два атома, изображены на рис. 4-11. Во всем диапазоне изменений q имеем $\omega_{\text{оп}} > \omega_{\text{ак}}$. Частота оптических колебаний незначительно уменьшается при увеличении q . Наибольшее отличие величины $\omega_{\text{оп}}$ от $\omega_{\text{ак}}$ имеет место при $q = 0$.

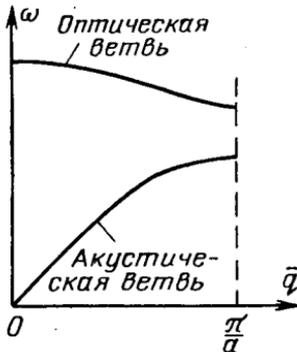


Рис. 4-11. Оптические и акустические ветви колебаний одномерной цепочки, в элементарной ячейке которой находятся два атома.

4-8. ТЕПЛОВЫЕ КОЛЕБАНИЯ АТОМОВ ТРЕХМЕРНОЙ РЕШЕТКИ. СТАТИСТИКА ФОНОНОВ

Рассмотренные простые случаи тепловых колебаний одномерной цепочки атомов позволяют высказать общие соображения о тепловом движении атомов трехмерного кристалла. В объемном кристалле сохраняются основные закономерности, имеющие место в одномерном случае. 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 . Колебание характеризуется длиной волны $\lambda = 2\pi/q$ и частотой ω_q . Волновой вектор q имеет столько разрешенных значений, сколько имеется в кристалле элементарных ячеек. При этом величины q меняются в пределах

$$-\pi \leq a_i q_i \leq \pi, \quad (4-127)$$

где 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 в кристалле будет $3sN$ волн.

В трехмерном кристалле имеют место два типа упругих волн: продольные — волны сжатий и разрежений и поперечные — волны деформаций сдвига. В первом случае 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 $2s$ типов поперечных волн и s типов продольных волн, всего $3s$ типов волн, т. е. должно быть $3s$ различных ветвей колебаний. Другими словами, каждому значению q соответствует $3s$ колебаний, различающихся частотой и поляризацией. Три ветви колебаний являются акустическими. Для них в области длинных волн имеет место $v_{\phi} = v_{гр} = v_{зв}$. 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 $v_{\phi} \rightarrow \infty$ при $q \rightarrow 0$.

В общем случае дисперсионные кривые для волн различной поляризации 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.

В случае германия и кремния в элементарной ячейке находятся два атома, поэтому они имеют оптические и акустические колебательные спектры, каждый из которых содержит одну продольную ветвь и две поперечные ветви колебаний, являющиеся вырожденными. На рис. 4-12 в качестве примера приведен колебательный спектр германия, для которого он был измерен экспериментально и также рас-

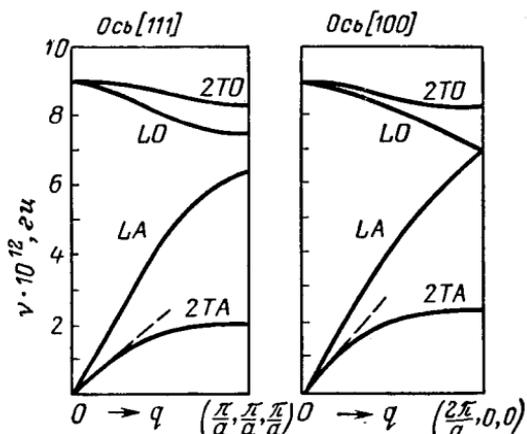


Рис. 4-12. Колебательный спектр решетки германия. TA и LA — поперечные и продольные акустические колебания, TO и LO — соответствующие оптические колебания.

считан теоретически. Зависимость частоты колебания ν от волнового вектора q на рисунке представлена для двух направлений вектора q : $[100]$ и $[111]$.

Как следует из рис. 4-12, частоты акустических колебаний стремятся 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 . При малых значениях волнового спектра q частота этих колебаний пропорциональна q для каждого направления распространения 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 $q \rightarrow 0$ ($\omega_{оп}(0) = \omega_0$). Она лежит в диапазоне 10^{13} гц . При $q \rightarrow 0$ фазовая скорость стремится к бесконечности, что является характерным для оптической ветви колебаний. С увеличением значения q частоты оптических колебаний решетки уменьшаются незначительно, а поэтому можно считать, что $\omega_{оп} = \omega_0$, в то время как частоты акустических колебаний резко возрастают. В области $q = \frac{\pi}{a}$, т. е. вблизи границы зоны Бриллюэна, различие между оптическими и акустическими колебаниями теряется.

Поскольку тепловые колебания решетки описываются $3sN$ нормальными колебаниями решетки, а энергия каждого нормального колебания квантованна, то полная колебательная энергия решетки может быть представлена в виде суммы энергий $3sN$ не взаимодействующих между собой гармонических осцилляторов:

$$E = \sum_{qj} E_{qj} = \sum_{qj} \hbar \omega_{qj} \left(n_{qj} + \frac{1}{2} \right) \quad (n_{qj} = 0, 1, 2 \dots), \quad (4-128)$$

где q — волновое число, имеющее N значений; j — ветвь колебания (всего имеется $3s$ ветвей колебаний или $3s$ типов колебаний); n_{qj} — главное квантовое число qj -го осциллятора, колеблющегося с частотой ω_{qj} .

Квантовый осциллятор с энергией $E_{qj} = \hbar \omega_{qj} \left(n_{qj} + \frac{1}{2} \right)$ может изменить энергию на величину $\Delta E_{qj} = \hbar \omega_{qj} \Delta n_{qj}$. При этом изменение квантового числа осциллятора согласно квантовой механике должно быть равно:

$$\Delta n_{qj} = \pm 1. \quad (4-129)$$

Квант энергии $\hbar \omega_{qj}$ 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 — фотон можно рассматривать как носитель кванта энергии колебаний решетки, т. е. рассматривать как свободную квазичастицу. Таким образом, колебания решетки в энергетическом отношении эквивалентны фотонному газу. В этом случае $\Delta n_{qj} = +1$ будет означать рождение фотона, а $\Delta n_{qj} = -1$ — уничтожение фотона.

Статистика фононов. Согласно уравнению (4-128) число фононов с энергией $\hbar\omega_{qj}$ может быть сколь угодно велико. Для того чтобы определить число фононов с энергией $\hbar\omega_{qj}$, существующих при данной температуре T , необходимо знание средней энергии осциллятора.

Пусть осциллятор в тепловом равновесии при температуре T находится в s -м квантовом состоянии с энергией E_s . Вероятность W_s того, что осциллятор находится в этом состоянии, согласно Больцману равна:

$$W_s = C e^{-\frac{E_s}{k_0 T}}. \quad (4-130)$$

Коэффициент C определяется из условия нормировки

$$\sum_s W_s = C \sum_s e^{-\frac{E_s}{k_0 T}} = 1, \quad (4-131)$$

откуда

$$C = \left(\sum_s e^{-\frac{E_s}{k_0 T}} \right)^{-1}. \quad (4-132)$$

Средняя энергия осциллятора, т. е. энергия при равновесном состоянии, равна сумме энергий E_s , помноженных на соответствующие вероятности W_s , т. е.

$$\langle E_s \rangle = \sum_{n=0}^{\infty} E_n W_n = \frac{\sum_{n=0}^{\infty} E_n e^{-\frac{E_n}{k_0 T}}}{\sum_{n=0}^{\infty} e^{-\frac{E_n}{k_0 T}}} = - \frac{d}{d(1/k_0 T)} \ln \sum_{n=0}^{\infty} e^{-\frac{E_n}{k_0 T}}. \quad (4-133)$$

Найдем значение суммы, учитывая энергию осциллятора $E_s = \hbar\omega_s (n + 1/2)$, где $n = 0, 1, 2, \dots$. Определить ее можно как сумму геометрической прогрессии:

$$\begin{aligned} \sum_{n=0}^{\infty} e^{-\frac{E_n}{k_0 T}} &= \sum_{n=0}^{\infty} e^{-\frac{\hbar\omega_s (n + 1/2)}{k_0 T}} = \\ &= e^{-\frac{\hbar\omega_s}{2k_0 T}} \left(1 + e^{-\frac{\hbar\omega_s}{k_0 T}} + e^{-\frac{2\hbar\omega_s}{k_0 T}} + \dots \right) = \frac{e^{-\frac{\hbar\omega_s}{2k_0 T}}}{1 - e^{-\frac{\hbar\omega_s}{k_0 T}}}. \end{aligned} \quad (4-134)$$

Если подставим выражение (4-134) в (4-133) и проведем дифференцирование по аргументу $1/k_0T$, то получим:

$$\begin{aligned} \langle E_s \rangle &= -\frac{d}{d(1/k_0T)} \ln \frac{e^{-\frac{\hbar\omega_s}{2k_0T}}}{1 - e^{-\frac{\hbar\omega_s}{k_0T}}} = \frac{\hbar\omega_s}{2} + \frac{d}{d(1/k_0T)} \ln \left(1 - e^{-\frac{\hbar\omega_s}{k_0T}} \right) = \\ &= \frac{\hbar\omega_s}{2} + \frac{\hbar\omega_s e^{-\frac{\hbar\omega_s}{k_0T}}}{1 - e^{-\frac{\hbar\omega_s}{k_0T}}} = \frac{\hbar\omega_s}{2} + \frac{\hbar\omega_s}{e^{\frac{\hbar\omega_s}{k_0T}} - 1}. \end{aligned} \quad (4-135)$$

Член $\frac{1}{2} \hbar\omega_s$, не зависящий от температуры, называется нулевой энергией осциллятора. Поскольку при взаимодействии с электронами участие принимают только те нормальные колебания, которые возбуждены выше нулевого энергетического уровня, то нулевую энергию можно не учитывать. Поэтому равновесное значение энергии гармонического осциллятора можно записать в виде

$$\langle E_s \rangle = \frac{\hbar\omega_s}{e^{\frac{\hbar\omega_s}{k_0T}} - 1}. \quad (4-136)$$

Обозначим через $\langle n \rangle$ отношение $\langle E_s \rangle$ и $\hbar\omega_s$:

$$\langle n \rangle = \frac{\langle E_s \rangle}{\hbar\omega_s} = \frac{1}{e^{\frac{\hbar\omega_s}{k_0T}} - 1}. \quad (4-137)$$

Значение $\langle n \rangle$ есть среднее число фононов (их концентрация) с энергией $\hbar\omega_s$ при тепловом равновесии в одном квантовом состоянии, т. е. в ячейке фазового пространства объемом h^3 с энергией $\hbar\omega_s$. Другими словами, $\langle n \rangle$ есть среднее число фононов в данном нормальном колебании (j q). Из выражения (4-136) следует, что фононы подчиняются статистике Бозе — Эйнштейна. Число фононов зависит от температуры. При высоких температурах, когда $T \gg \hbar\omega_s/k_0$, число $\langle n \rangle$ фононов в одной ячейке фазового пространства равно:

$$\langle n \rangle = \frac{k_0T}{\hbar\omega_s}, \quad (4-138)$$

а средняя энергия осциллятора равна:

$$\langle E_s \rangle = \hbar\omega_s \langle n \rangle = k_0T. \quad (4-139)$$

При низких температурах, когда $T \ll \hbar\omega_s$,

$$\langle n \rangle \approx e^{-\frac{\hbar\omega_s}{k_0T}}. \quad (4-140)$$

Кроме того, при взаимодействии с электронами или друг с другом фононы возникают и исчезают, следовательно, фононы ведут себя не как газ обычных частиц, а как квазичастицы.

Проведем подсчет полного числа фононов в кристалле единичного объема. Пусть в интервале энергий от $\hbar\omega$ до $\hbar(\omega + d\omega)$ имеется dN_ϕ фононов. Это число фононов определяется числом квантовых состояний фононов dZ и средним числом $\langle n \rangle$ фононов, приходящихся на одно квантовое состояние. Если $\omega = v_{0зв}q$, где $v_{0зв}$ — средняя скорость звука акустических ветвей, а квазиимпульс фонона $p = \hbar q = \hbar\omega/v_{0зв}$, то число квантовых состояний, т. е. ячеек объемом h^3 , на интервал dp с учетом трех возможных поляризаций фонона (одну продольную и две поперечные) будет равно:

$$dZ = 3 \frac{4\pi p^2 dp}{h^3} = \frac{3\omega^2 d\omega}{2\pi^2 v_{0зв}^3}. \quad (4-141)$$

Тогда на интервал энергии $\hbar d\omega$ будет приходиться фононов в количестве

$$dN_\phi = \langle n \rangle dZ = \frac{3}{2\pi^2 v_{0зв}^3} \frac{\omega^2 d\omega}{e^{\frac{\hbar\omega}{k_0 T}} - 1}. \quad (4-142)$$

Полное число фононов в кристалле единичного объема равно

$$N_\phi = \frac{3}{2\pi^2 v_{0зв}^3} \int_0^{\omega_{\max}} \frac{\omega^2 d\omega}{e^{\frac{\hbar\omega}{k_0 T}} - 1} = \frac{3}{2\pi^2 v_{0зв}^3} \left(\frac{k_0 T}{\hbar}\right)^3 \int_0^{\theta/T} \frac{x^2 dx}{e^x - 1}, \quad (4-143)$$

где ω_{\max} — максимальная частота продольных акустических колебаний;

$$x = \frac{\hbar\omega}{k_0 T},$$

а

$$\theta = \frac{\hbar\omega_{\max}}{k_0} \quad (4-144)$$

есть характеристическая температура твердого тела, или температура Дебая. Так как $\omega_{\max} \approx \frac{v_{зв}}{a} \approx 10^{13} \text{ сек}^{-1}$, $\hbar = 1,05 \cdot 10^{-27} \text{ эрг/сек}$, $k_0 = 1,38 \cdot 10^{-16} \text{ эрг/град}$, то $\theta \approx 100^\circ \text{ К}$. Для большинства твердых тел θ лежит в пределах $100-300^\circ \text{ К}$.

При низких температурах, когда $T \ll \theta$, в выражении (4-143) верхний предел можно заменить на ∞ и число фононов $N_\phi \sim T^3$. При высоких температурах, когда $T \gg \theta$, т. е. когда $\theta/T \ll 1$ и $x \ll 1$, а потому и $e^x = x + 1$, число фононов $N_\phi \sim T$. Зависимость концентрации фононов от температуры представлена на рис. 4-13.

Как следует из выражений (4-115) и (4-116), а также рис. 4-11 и 4-12, при низких температурах возбуждаются только длинноволновые колебания акустической ветви с малой энергией $\hbar\omega_{а.к} =$

$= \hbar \nu_{\text{оп}} q$. При более высоких температурах возбуждаются все акустические колебания. Оптические колебания возбуждаются при сравнительно высоких температурах, так как частоты оптических колебаний больше частот акустических колебаний. В связи с этим можно говорить об оптических и акустических фононах.



Рис. 4-13. Зависимость концентрации фононов от температуры.

Частоты оптических колебаний мало меняются во всем диапазоне изменений волнового числа q . Поэтому можно считать, что $\omega_{q\text{оп}} = \omega_0$. Введем температуру Дебая, соответствующую предельной частоте данной оптической ветви, как

$$\theta_{0j} = \frac{\hbar \omega_{0j}}{k_0}.$$

При $T \gg \theta_0$ в кристалле будет большое число оптических фононов, а при $T \ll \theta_0$ оптические фононы практически не возбуждаются. Для большинства веществ θ_0 лежит в пределах $100\text{--}800^\circ \text{K}$. В табл. 4-1 приведены значения энергии продольных оптических фононов $\hbar \omega_0$ и соответствующие ей температуры Дебая.

Таблица 4-1

Ширина запрещенной зоны, максимальная частота продольных оптических фононов ω_0 , их энергия $\hbar \omega_0$ и температура Дебая для некоторых полупроводников

Полупроводники	Ширина запрещенной зоны (300°K), эВ	$\omega_0 \cdot 10^{-13}$, эц	$\hbar \omega_0$, эВ	θ_0 , °K
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

4-9. 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 фонон с энергией $\hbar \omega_q$ и квазиимпульсом $p = \hbar q$, и число фононов возрастет на единицу. $p = \hbar q$ есть импульс, который приобретает решетка вследствие соударения с электронами. 2) Электрон при взаимодействии получает

часть энергии от решетки, так что квантовое число определенного колебания с частотой ω_q уменьшается на единицу. В этом случае исчезает фонон с энергией $\hbar\omega_q$ и квазиимпульсом $p = \hbar q$ и 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 называется *однофононным*. Так как число фононов определяется температурой, то и расстояние электронов на тепловых колебаниях решетки должно зависеть от температуры. В кристалле возможен и *многочононый* механизм рассеяния, при котором имеет место поглощение или испускание более чем одного фонона. Однако вероятность такого процесса мала по сравнению с вероятностью однофононного рассеяния.

Пусть до столкновения электрон имел волновой вектор k и энергию $E(k)$, а после столкновения его волновой вектор стал k' и энергия $E'(k')$. Фонон, возникающий или исчезающий при столкновении, обладает энергией $\hbar\omega_q$ и волновым вектором q . При взаимодействии электрона с фононом будут выполняться два закона: закон сохранения энергии и закон сохранения квазиимпульса. Поэтому при поглощении электроном фонона

$$E'(k') = E(k) + \hbar\omega_q;$$

$$k' = k + q,$$

а число фононов в конечном состоянии уменьшится на единицу, т. е.

$$N'_q = N_q - 1.$$

При испускании электроном фонона

$$E'(k') = E(k) - \hbar\omega_q,$$

$$k' = k - q, \tag{4-145}$$

при этом

$$N'_q = N_q + 1. \tag{4-146}$$

Рассеяние носителя заряда может быть упругим и неупругим. Для характеристики рассеяния необходимо знание относительного изменения энергии δ электрона за одно столкновение

$$\delta = \frac{E' - E}{E} = \frac{\Delta E}{E} \tag{4-147}$$

или за единицу времени

$$A = \frac{\Delta E}{\tau} = \frac{E\delta}{\tau}. \tag{4-148}$$

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 сингонии, энергетическая схема которого представлена на рис. 4-14, где a — постоянная решетки в отсутствие деформации. Пусть зона проводимости в отсутствие такого кристалла простая, так что энергия электрона равна:

$$E = \frac{\hbar^2 \kappa^2}{2m^*}. \quad (4-149)$$

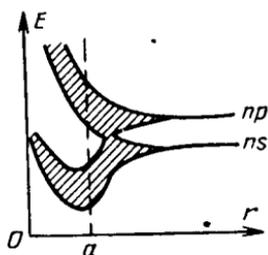


Рис. 4-14. Энергетическая схема атомного полупроводника.

4-14, при сжатии кристалла, сопровождающемся уменьшением 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 , согласно рис. 4-14 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. 4-15.



Рис. 4-15. Изменение энергии дна зоны проводимости и потолка валентной зоны атомного полупроводника под воздействием продольных акустических колебаний решетки.

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 фононах. Закон сохранения энергии согласно (4-149) запишется в виде

$$\frac{\hbar^2 \kappa^2}{2m^*} = \frac{\hbar^2 (\kappa \pm q)^2}{2m^*} = \frac{\hbar^2 \kappa^2}{2m^*} \pm \hbar \omega q, \quad (4-150)$$

откуда

$$q^2 \pm 2\kappa q \cos \vartheta \mp \frac{2m^*}{\hbar} \omega q = 0, \quad (4-151)$$

где ϑ — угол между направлениями векторов κ и q .

Если электрон взаимодействует только с длинноволновыми фононами, то $\omega q = v_{зв} q$, где $v_{зв}$ — скорость распространения продольной звуковой волны. Учитывая это, из уравнения (4-151) находим:

$$q = \mp 2\kappa \cos \vartheta \pm \frac{2m^* v_{зв}}{\hbar}. \quad (4-152)$$

Так как $\frac{\hbar^2 \kappa^2}{2m^*} = \langle E \rangle = k_0 T$, то $\kappa = \frac{1}{\hbar} \sqrt{2m^* k_0 T}$, поэтому

$$\frac{m^* v_{зв}}{\hbar \kappa} = \sqrt{\frac{m^* v_{зв}^2}{2k_0 T}} = \sqrt{\frac{T_{кр}}{T}}, \quad (4-153)$$

где $T_{кр} = \sqrt{\frac{m^* v_{зв}^2}{2k_0}}$. Если принять скорость звука $v_{зв} = 3 \cdot 10^5$ см/сек, $m^* = m = 10^{-27}$ г, а $k_0 = 10^{-16}$ эрг/град, то $T_{кр} = 1^\circ$ К. Следовательно, при температурах, намного превышающих 1° К, можно пренебречь вторым слагаемым в выражении (4-152) по сравнению с первым и записать

$$q = \pm 2\kappa \cos \vartheta. \quad (4-154)$$

В зависимости от ϑ волновое число q может меняться от 0 до 2κ , т. е. в среднем на κ . Поскольку при $T = 300^\circ$ К $\kappa \approx 10^7$ см $^{-1}$, то q может принимать значения от 0 до $2 \cdot 10^7$ см $^{-1}$, что соответствует энергии фононов от 0 до $\hbar \omega_q = \hbar v_{зв} \kappa = 10^{-27} \cdot 10^5 \cdot 10^7 = 10^{-15}$ эрг = $6 \cdot 10^{-4}$ эв. Отсюда следует, что электрон рассеивается на длинноволновых фононах, поглощая или излучая фонон с $q \approx \kappa$. Так как $E = \frac{\hbar^2 \kappa^2}{2m^*} = 3 \cdot 10^{-2}$ эв, то

$$\frac{\hbar \omega_q}{E} = \frac{\hbar \omega_q}{k_0 T} \ll 1. \quad (4-155)$$

Следовательно, рассеяние электронов на продольных акустических колебаниях является упругим, и основную роль в рассеянии играют длинноволновые фононы с энергией, значительно меньшей энергии электронов. Поэтому $\kappa' = \kappa$ и $E' = E$, поскольку в равенстве (4-150) можно не учитывать энергию фонона.

При низких температурах средняя скорость движения электрона приближается к скорости распространения упругих волн в кристалле и энергией фонона в уравнении (4-152) 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 (4-39):

$$\begin{aligned} \frac{1}{\tau} &= \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left[1 - \frac{\kappa'_\chi}{\kappa_\chi} \right] d\tau_{\mathbf{k}'} = \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left(\mp \frac{q_\chi}{\kappa_\chi} \right) d\tau_{\mathbf{k}'} = \\ &= -\frac{1}{4\pi^3} \int_{(V_B)} W^+(\mathbf{k}, \mathbf{k}') \frac{q_\chi}{\kappa_\chi} d\tau_{\mathbf{k}'} + \frac{1}{4\pi^3} \int_{(V_B)} W^-(\mathbf{k}, \mathbf{k}') \frac{q_\chi}{\kappa_\chi} d\tau_{\mathbf{k}'} \end{aligned} \quad (4-156)$$

Здесь первый член правой части учитывает поглощение, а второй — испускание фононов при рассеянии электронов.

Вероятность перехода электрона из состояния \mathbf{k} в состояние \mathbf{k}' , как показывает расчет¹, равна в случае поглощения фонона

$$W^+(\mathbf{k}, \mathbf{k}') = \frac{4\pi}{9} \frac{C^2 q^2}{NM\omega_q} N_q \delta(E' - E - \hbar\omega_q), \quad (4-157)$$

¹ Ансельм А. И., Введение в теорию полупроводников, Физматгиз, 1962, стр. 278.

а в случае испускания фонона

$$W^-(\kappa, \kappa') = \frac{4\pi}{9} \frac{C^2 q^2}{NM\omega_q} (N_q + 1) \delta(E' - E - \hbar\omega_q), \quad (4-158)$$

где $\delta(E' - E - \hbar\omega_q)$ — дельта-функция; C — постоянная Блоха, приблизительно пропорциональная дебаевской температуре кристалла; N — число элементарных ячеек в кристалле; M — масса атома.

В элементарных процессах переноса функция распределения фононов мало отличается от равновесной $\langle N_q \rangle_{\text{равн}}$, поэтому можно считать, что

$$N_q = \langle \bar{N}_q \rangle_{\text{равн}} = \frac{1}{\frac{\hbar\omega_q}{e^{k_0 T}} - 1}. \quad (4-159)$$

Но $\frac{\hbar\omega_q}{k_0 T} \approx \frac{\hbar v_{\text{зв}} \kappa}{k_0 T} = \sqrt{\frac{T_{\text{кр}}}{T}} \ll 1$, поэтому формулу (4-159) можно разложить в ряд

$$\langle N_q \rangle_{\text{равн}} = \langle N_{q+1} + 1 \rangle_{\text{равн}} = \frac{k_0 T}{\hbar v_{\text{зв}} q}. \quad (4-160)$$

Если в выражение (4-156) подставить (4-157) и (4-158), учтя равенства (4-159) и (4-160), затем перейти к сферической системе координат, то в результате получим, что

$$\tau = \frac{9\pi}{4\sqrt{2}} \frac{\hbar^4 M v_{\text{зв}}^2}{C^2 a^3 k_0 m^{*3/2}} \frac{1}{T E^{1/2}}, \quad (4-161)$$

где a — параметр решетки.

Таким образом, время релаксации при рассеянии электронов на акустических колебаниях решетки атомных полупроводников имеет зависимость от энергии вида

$$\tau(\kappa) = \frac{\tau_0}{m^{*3/2} T} E^{-1/2}, \quad (4-162)$$

где

$$\tau_0 = \frac{9\pi}{4\sqrt{2}} \frac{\hbar^4 M v_{\text{зв}}^2}{C^2 a^3 k_0},$$

а длина свободного пробега

$$l = v\tau = \frac{9\pi}{4} \frac{\hbar^4 M v_{\text{зв}}^2}{C^2 a^3 k_0 m^{*2}} \frac{1}{T} = \frac{\sqrt{2} \tau_0}{m^{*2}} T^{-1} \quad (4-163)$$

обратно пропорциональна температуре и 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 $\mu_n = 0,39 \text{ м}^2/\text{в} \cdot \text{сек}$. Если принять $m^* = 0,3 m$, то $e/m^* = 6 \cdot 10^{11} \text{ к/кг}$. Следовательно,

$$\tau = \frac{m^* \mu_n}{e} = \frac{0,39}{6 \cdot 10^{11}} = 6,5 \cdot 10^{-13} \text{ сек}.$$

Из сравнения формул (4-162) и (4-68) видно, что при повышении энергии носителей заряда время релаксации при рассеянии на ионизированной 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 значительно сказываться при высоких температурах.

В полярных полупроводниках, а также в полупроводниковых соединениях типа $A^{III}B^V$, 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 разноименные 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 значительно сильнее рассеивают, чем поперечные. Каждое столкновение электрона с оптическим фононом приводит к возникновению или исчезновению фонона с энергией $\hbar\omega_{оп}$. Поскольку частота оптического фонона слабо зависит от квазимпульса q , то при соударении энергия электрона либо увеличивается, либо уменьшается на одинаковую величину, равную

$$\pm \hbar\omega_{оп} = \pm \hbar\omega_0, \quad (4-164)$$

где ω_0 — предельная частота продольной оптической ветви.

Энергия, передаваемая носителем заряда решетке за одно столкновение, определяется величиной энергии возникшего фонона, умноженной на отношение вероятности испускания минус вероятность поглощения к сумме этих вероятностей. Но вероятность поглощения фонона согласно формуле (4-157) пропорциональна концентрации фононов N_q , а вероятность испускания фонона согласно формуле (4-158) пропорциональна $N_q + 1$. Следовательно, энергия, передаваемая электроном фонону за одно столкновение, равна:

$$\begin{aligned} \Delta E &= \hbar\omega_0 \frac{W^-(\kappa, \kappa') - W^+(\kappa, \kappa')}{W^-(\kappa, \kappa') + W^+(\kappa, \kappa')} = \hbar\omega_0 \frac{(N_q + 1) - N_q}{(N_q + 1) + N_q} = \\ &= \hbar\omega_0 \frac{1}{2N_q + 1} = \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2k_0T}. \end{aligned} \quad (4-165)$$

При высоких температурах, когда $T \gg \theta_0$ или $k_0T \gg \hbar\omega_0$, т. е. когда энергия фонона $\hbar\omega_0$ много меньше энергии электрона k_0T , тогда

$$\delta_{оп} = \frac{\Delta E}{E} = \frac{(\hbar\omega_0)^2}{2(k_0T)^2} \ll 1 \quad (4-166)$$

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 (4-156). Для случая рассеяния на оптических фононах, как показывает расчет¹, время релаксации

$$\tau_{\text{оп}} = \frac{\sqrt{2}}{4\pi} \frac{Ma^3 (\hbar\omega_0)^2 E^{1/2}}{Z^2 e^4 m^{*1/2} k_0 T}, \quad (4-167)$$

а длина свободного пробега

$$l_{\text{оп}} = \tau_{\text{оп}} v = \frac{aM}{2\pi m^*} \left(\frac{\hbar\omega_0}{Ze^2/a} \right)^2 \frac{E}{k_0 T}, \quad (4-168)$$

где Ze — заряд иона; $M = \frac{m_1 m_2}{m_1 + m_2}$ — приведенная масса ионов; a — расстояние между ближайшими разноименными ионами.

При низких температурах, когда $T \ll \theta_0$ или $k_0 T \ll \hbar\omega_0$, большинство электронов, у которых энергия $E \ll \hbar\omega_0$, могут только поглощать оптические фононы. В результате такого поглощения фононов электрон переходит в интервал энергии от $\hbar\omega_0$ до $2\hbar\omega_0$, т. е. будет очень сильно изменять свою энергию. Поскольку вероятность испускания фонона (4-158) в $(N_q + 1)/N_q$ раз превышает вероятность поглощения фонона (4-157), время пребывания электрона в состоянии с большой энергией будет очень мало. Можно считать, что электрон сразу после поглощения фонона испускает фонон той же энергии. Фактически рассеяние можно рассматривать как процесс обмена фононами, поэтому энергия электрона практически остается неизменной, но направление его квазиимпульса сильно изменится. Это позволяет считать, что и при низких температурах, значительно меньших характеристической $\theta_0 = \hbar\omega_0/k_0$, электрон, взаимодействуя с оптическими колебаниями решетки, рассеивается упруго.

Если учесть, что в случае низких температур

$$N_q = \langle N_q \rangle_{\text{равн}} = \frac{1}{\frac{\hbar\omega_0}{e^{k_0 T}} - 1} = e^{-\frac{\hbar\omega_0}{k_0 T}}, \quad (4-169)$$

расчет для времени релаксации по формуле (4-156), в которой учитывается только поглощение фонона, дает выражение вида

$$\tau_{\text{оп}} = \frac{3\sqrt{2}}{4\pi} \frac{Ma^3 (\hbar\omega_0)^{3/2}}{Z^2 e^4 m^{*1/2}} e^{\frac{\hbar\omega_0}{k_0 T}}, \quad (4-170)$$

которое не зависит от энергии электрона, но экспоненциально зависит от температуры. Длина свободного пробега в случае рассея-

¹ Ансельм А. И., Введение в теорию полупроводников, Физматгиз, 1962.

ния на оптических фонах при низких температурах определяется соотношением вида

$$l_{\text{оп}} = \frac{3aM}{2\pi m^*} \left(\frac{\hbar\omega_0}{Ze^2/a} \right)^2 e^{\frac{\hbar\omega_0}{k_0T}} \left(\frac{E}{k_0T} \right)^{1/2}. \quad (4-171)$$

Г л а в а п я т а я

КИНЕТИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

5-1. ЭЛЕКТРОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ

Определим удельную проводимость однородного невырожденного полупроводника, изоэнергетические поверхности которого представляют сферы. Пусть полупроводник находится в постоянном однородном электрическом поле напряженностью \mathcal{E} .

В элементе объема $d\tau_{\mathbf{k}}$ кристалла единичного объема согласно выражению (4-2) количество электронов равно:

$$dn = f(\mathbf{k}) \frac{d\tau_{\mathbf{k}}}{4\pi^3}.$$

Эти электроны, движущиеся под действием электрического поля со скоростью \mathbf{v} , создадут элементарную плотность тока

$$dj = -ef(\mathbf{k}) \frac{d\tau_{\mathbf{k}}}{4\pi^3} \mathbf{v}. \quad (5-1)$$

Если в полупроводнике имеются два сорта носителей заряда, то полная плотность тока будет равна:

$$\mathbf{j} = \mathbf{j}_n + \mathbf{j}_p = -\frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_n f_n(\mathbf{k}) d\tau_{\mathbf{k}} + \frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_p f_p(\mathbf{k}') d\tau_{\mathbf{k}'}, \quad (5-2)$$

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 выражение (5-2) подставить f_0 , то $\mathbf{j} = 0$; это означает, что при термодинамическом равновесии в веществе нет тока.

С учетом этого, если неравновесная функция $f(\mathbf{k}) = f_0 + f_1(\mathbf{k})$, вместо выражения (5-2) получим:

$$\mathbf{j} = -\frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_n f_1^{(n)}(\mathbf{k}) d\tau_{\mathbf{k}} + \frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_p f_1^{(p)}(\mathbf{k}') d\tau_{\mathbf{k}'}. \quad (5-3)$$

Неравновесная функция распределения может быть рассчитана из условия стационарного состояния

$$-\left(\frac{\partial f}{\partial t} \right)_{\text{поп}} = \left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = -\frac{f_1(\mathbf{k})}{\tau(\mathbf{k})}. \quad (5-4)$$

На электрон с энергией $E = \hbar^2 \kappa^2 / 2m_n^*$, находящейся во внешнем электрическом поле, действует сила $F = \hbar \frac{d\kappa}{dt} = -e\mathcal{E}$. Поскольку полупроводник однородный, то

$$\nabla_{\tau} f = 0 \quad (5-5)$$

и изменение функции распределения под влиянием поля будет определяться выражением следующего вида:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{пол}} = \frac{\partial f_n}{\partial E} \frac{\partial E}{\partial \kappa} \frac{\partial \kappa}{\partial t} = -\frac{e\hbar}{m_n^*} \frac{\partial f_n}{\partial E} (\kappa \mathcal{E}). \quad (5-6)$$

Из соотношения (5-4) находим $f_n^{(1)}(\kappa)$, используя (5-6), и подставляем его в (5-3). В результате получим выражение для плотности электронного тока

$$\mathbf{j}_n = -\frac{e^2 \hbar^2}{4\pi^3 m_n^{*2}} \int_{(V_B)} \tau_e \frac{\partial f_n}{\partial E} (\kappa \mathcal{E}) \kappa d\tau_{\kappa}. \quad (5-7)$$

Здесь τ_e — время релаксации электрона.

Будем считать, что неравновесная функция $f_n(\kappa)$ мало отличается от равновесной f_0 , тогда в выражении (5-7) можно заменить $f_n(\kappa)$ на f_0 . Для невырожденного полупроводника, у которого отсчет энергии производится от дна зоны проводимости, количество электронов в зоне проводимости в соответствии с формулой (3-60) определяется соотношением вида

$$n = N_c e^{-\frac{F_c - F}{k_0 T}} = N_c e^{-\frac{F}{k_0 T}}. \quad (5-8)$$

Учитывая это, находим:

$$\frac{\partial f_0}{\partial E} = -\frac{1}{k_0 T} e^{-\frac{E - F}{k_0 T}} = -\frac{n}{N_c k_0 T} e^{-\frac{E}{k_0 T}}. \quad (5-9)$$

Подставляя уравнение (5-9) в (5-7), получаем:

$$\mathbf{j}_n = \frac{e^2 \hbar^2 n}{4\pi^3 m_n^{*2} k_0 T N_c} \int_{(V_B)} \tau_e e^{-\frac{E}{k_0 T}} (\kappa \mathcal{E}) \kappa d\tau_{\kappa}. \quad (5-10)$$

Если воспользоваться соотношением¹

$$\int F(\kappa) (\kappa \mathbf{a}) \kappa d\tau_{\kappa} = \frac{4\pi}{3} \mathbf{a} \int_0^{\infty} F(\kappa) \kappa^4 d\kappa, \quad (5-11)$$

то выражение (5-10) запишется в виде

$$\mathbf{j}_n = \frac{e^2 \hbar^2 n \mathcal{E}}{4\pi^3 m_n^{*2} N_c k_0 T} \frac{4\pi}{3} \int_0^{\infty} \tau_e e^{-\frac{E}{k_0 T}} \kappa^4 d\kappa. \quad (5-12)$$

¹ Ансельм А. И., Введение в теорию полупроводников. Физматгиз, 1962, стр. 416.

Полагая $\alpha = E/k_0T$, получаем:

$$\kappa^4 dk = \frac{(2m_n^*k_0T)^{3/2}}{2h^5} \alpha^{3/2} d\alpha, \quad (5-13)$$

а подставляя выражение (5-13) в (5-12) и учитывая, что $N_c = 2 \left(\frac{2\pi m_n^*k_0T}{h^2} \right)^{3/2}$, получаем:

$$\mathbf{j}_n = \frac{e^2 n}{m_n^*} \mathcal{E} \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha. \quad (5-14)$$

Введем обозначение

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha. \quad (5-15)$$

Тогда

$$\mathbf{j}_n = \frac{e^2 n \langle \tau_e \rangle}{m_n^*} \mathcal{E} = en\mu_n \mathcal{E} = \sigma_n \mathcal{E}, \quad (5-16)$$

где $\mu_n = \frac{e \langle \tau_e \rangle}{m_n^*}$ — подвижность электронов; (5-17)

$$\sigma_n = en\mu_n \quad (5-18)$$

есть электропроводность, обусловленная электронами.

Соответственно дырочная составляющая тока будет равна:

$$\mathbf{j}_p = \frac{e^2 p \langle \tau_h \rangle}{m_p^*} \mathcal{E} = ep\mu_p \mathcal{E} = \sigma_p \mathcal{E}. \quad (5-19)$$

Здесь среднее время релаксации дырок

$$\langle \tau_h \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_h e^{-\alpha} \alpha^{3/2} d\alpha. \quad (5-20)$$

Подвижность дырок

$$\mu_p = \frac{e \langle \tau_h \rangle}{m_p^*}. \quad (5-21)$$

Электропроводность, обусловленная дырками,

$$\sigma_p = ep\mu_p. \quad (5-22)$$

Следовательно, полная плотность тока полупроводника, имеющего два типа носителей заряда, равна:

$$\mathbf{j} = \mathbf{j}_n + \mathbf{j}_p = (\sigma_n + \sigma_p) \mathcal{E} = \sigma \mathcal{E}. \quad (5-23)$$

5.2. ЗАВИСИМОСТЬ ПОДВИЖНОСТИ НОСИТЕЛЕЙ ЗАРЯДА ОТ ТЕМПЕРАТУРЫ

Проведенный расчет показывает, что для определения подвижности носителей заряда необходимо знание среднего времени релаксации. Но на времени релаксации существенным образом сказыва-

вается механизм рассеяния, поэтому при столкновении носителей заряда с разными по природе дефектами кристаллической решетки подвижность по-разному будет зависеть от температуры.

При рассеянии носителей заряда в атомных полупроводниках на акустических колебаниях решетки время релаксации согласно формуле (4-162) равно:

$$\tau(\kappa) = \frac{\tau_0}{m_n^{*3/2}} T^{-1} E^{-1/2}, \quad (5-24)$$

поэтому

$$\mu(T) = \frac{4e}{3\sqrt{\pi}} \frac{\tau_0}{k_0^{1/2} m_n^{*3/2}} T^{-3/2}. \quad (5-25)$$

Подвижность можно выразить через длину свободного пробега. Используя формулу (4-163), находим:

$$\mu(T) = \frac{4el}{3(2\pi m_n^* k_0 T)^{1/2}}. \quad (5-26)$$

Из формулы (5-25) следует, что подвижность, обусловленная тепловыми колебаниями решетки, уменьшается с ростом температуры. Кроме того, величина подвижности обратно пропорциональна эф-

Таблица 5-1

Экспериментальные данные по подвижности электронов и дырок некоторых полупроводников

Полупроводник	Подвижность электронов, см ² /в-сек		Подвижность дырок, см ² /в-сек		Температурная зависимость решеточной подвижности		$b = \frac{\mu_n}{\mu_p}$	
	300° К	77° К	300° К	77° К	электроны	дырки	300° К	77° К
Ge	3 800	37 100	1 820	43 700	1,6	2,3	2,1	0,7
Si	1 300	45 500	500	11 600	2,6	2,3	2,6	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	46 000	24 000	150	1 200	2,0	2,4	30	20
GaSb	4 000	6 000	1 400	3 600	2,0	0,9	3	1,7
GaAs	8 500	21 000	420	4 200	1,0	2,1	20	5
GaP	110	500	75	420	1,5	1,5	1,5	1
AlSb	200	—	420	3 700	—	1,8	—	—

фективной массе носителей заряда в степени 5/2. Действительно, как показывает опыт, подвижность электронов больше подвижности дырок, а у полупроводников с малой эффективной массой носителей подвижность очень большая (см. табл. 5-1).

При рассеянии на ионах примеси время релаксации в зависимости от энергии имеет вид:

$$\tau_I = \frac{\sqrt{2} \bar{\epsilon}^2 m_n^{*1/2} E^{3/2}}{\pi Z^2 e^4 N_I \ln \left[1 + \left(\frac{\epsilon E}{Ze^2 N_I^{1/3}} \right)^2 \right]} \quad (5-27)$$

При усреднении этого выражения логарифмический член, являющийся медленно меняющейся функцией, можно вынести за знак интеграла, если в этот член подставить такое значение энергии, при

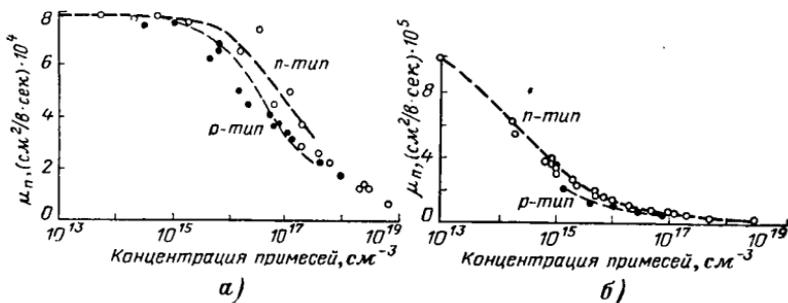


Рис. 5-1. Зависимость дрейфовой подвижности электронов от концентрации примесей в образцах антимида индия *n*- и *p*-типа при 300° К (а) и 77° К (б). Пунктирные кривые — рассчитанные теоретически.

котором остальная часть подинтегрального выражения достигает максимума. Для этого необходимо положить $E = 3k_0T$. Поэтому среднее время релаксации электрона будет равно:

$$\langle \tau_e \rangle = \frac{8\sqrt{2} \bar{\epsilon}^2 k_0^{3/2} m_n^{*1/2} T^{3/2}}{\pi^{3/2} Z^2 e^4 N_I \ln \left[1 + \left(\frac{3\epsilon k_0 T}{Ze^2 N_I^{1/3}} \right)^2 \right]} \quad (5-28)$$

а подвижность при рассеянии на ионах примеси запишется в виде

$$\mu_I(T) = \frac{8\sqrt{2} \bar{\epsilon}^2 k_0^{3/2} T^{3/2}}{\pi^{3/2} Z^2 e^4 m_n^{*1/2} N_I \ln \left[1 + \left(\frac{3\epsilon k_0 T}{Ze^2 N_I^{1/3}} \right)^2 \right]} = \frac{b \cdot \text{cm} \cdot \text{cm}^2}{\text{K} \cdot \text{cm}^{-3} \cdot \text{cm}^2} \quad (5-29)$$

Если пренебречь логарифмической зависимостью подвижности от температуры, что можно сделать при достаточно больших температурах, то можно считать, что

$$\mu_I \approx \mu_{0I} T^{3/2} \quad \frac{1}{\text{K} \cdot \text{cm}^{-3} \cdot \text{cm}^2} \quad (5-30)$$

Из формулы (5-30) следует, что подвижность электронов, обусловленная рассеянием на ионах примеси, уменьшается при уменьшении температуры. Это происходит потому, что с увеличением температуры кристалла возрастает тепловая скорость носителей заряда. Это приводит к ослаблению взаимодействия носителей с ионизированными

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. 5-1 представлена зависимость подвижности электронов от концентрации примесей в образцах антимонида индия n - и p -типа при 300 и 77° К.

При рассеянии носителей заряда на нейтральной примеси время релаксации (4-71) не зависит от температуры и энергии носителя. В силу этого *подвижность носителей, обусловленная рассеянием на нейтральных центрах, в явном виде не зависит от температуры и имеет вид:*

$$\mu_A = \frac{e^3 m_n^*}{20 \epsilon \hbar^3} \frac{1}{N_A}. \quad (5-31)$$

При рассеянии носителей на дислокациях согласно (4-72) $\tau_D \sim E^{-1/2}$ и подвижность

$$\mu_D(T) = \frac{e T^{-1/2}}{2^{3/2} \pi^{1/2} m_n^*{}^{1/2} R N_D}. \quad (5-32)$$

В реальных полупроводниках имеют место все виды нарушения строгой периодичности поля решетки, поэтому механизм рассеяния носит сложный характер. Если считать, что все виды механизма рассеяния независимы, то полная вероятность рассеяния равна сумме вероятностей рассеяния на каждом типе рассеивающих центров, т. е.

$$W(\mathbf{k}, \mathbf{k}') = \sum_i W_i(\mathbf{k}, \mathbf{k}'). \quad (5-33)$$

В этом случае полное время релаксации τ согласно выражению (4-39) будет равно:

$$\tau^{-1} = \frac{1}{4\pi^3} \int \sum_i W_i(\mathbf{k}, \mathbf{k}') (1 - \cos \theta) d\tau_{\mathbf{k}'} = \sum_i \tau_i^{-1}. \quad (5-34)$$

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 поля зависит от того, какой из механизмов рассеяния является преобладающим в данном интервале температур. При очень низких температурах в выражении (5-33) можно учитывать только рассеяние на нейтральной 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 подвижность носителей в зависимости от температуры согласно форму-

лам (5-25) и (5-30) будет определяться соотношением вида

$$\frac{1}{\mu} = \frac{m^*}{e} \sum_i \frac{1}{\tau_i} = aT^{3/2} + bT^{-3/2}, \quad (5-35)$$

где a и b — постоянные величины.

Как следует из формулы (5-35), в атомных полупроводниках с ростом температуры подвижность растет пропорционально $T^{3/2}$, если рассеяние происходит только на ионах примеси, затем проходит через максимум и уменьшается пропорционально $T^{-3/2}$, если рассеивающими центрами являются только акустические фононы.

Как показывает эксперимент, у полупроводников действительно наблюдается такой ход зависимости подвижности от температуры. В качестве примера на рис. 5-2 даны температурные зависимости подвижностей электронов и дырок в индия.

Но, как следует из данных табл. 5-1, у большинства полупроводников показатель степени p температурной зависимости подвижности не соответствует теоретическому значению

$p = -3/2$, обусловленному рассеянием на акустических колебаниях решетки. Это объясняется тем, что в реальных полупроводниках при данной температуре имеют место и другие механизмы рассеяния, такие как рассеяние на оптических колебаниях, двухфононное рассеяние, рассеяние на носителях заряда.

Учитывая зависимость подвижности и концентрации носителей заряда от температуры, электропроводность собственного полупроводника можно записать в виде

$$\sigma = en_i \mu_n + ep_i \mu_p = AT^{3/2+p} e^{-\frac{\Delta E_g}{2k_0 T}}. \quad (5-36)$$

Множитель $T^{3/2+p}$ медленно меняется с температурой, тогда как множитель $e^{-\frac{\Delta E_g}{2k_0 T}}$ сильно зависит от температуры, если $\Delta E_g \gg k_0 T$. Следовательно, для не слишком высоких температур можно

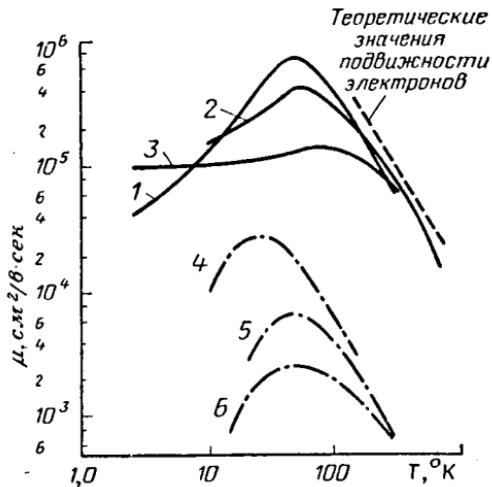


Рис. 5-2. Температурные зависимости подвижности электронов и дырок в InSb. 1 — $n = 2 \cdot 10^{14}$; 2 — $n = 10^{15}$; 3 — $n = 10^{16}$; 4 — $p = 3 \cdot 10^{14}$; 5 — $p = 4 \cdot 10^{15}$ и 6 — $p = 3 \cdot 10^{16} \text{ см}^{-3}$.

считать, что

$$AT^{3/2+p} = \sigma_0 = \text{const}, \quad (5-37)$$

и выражение (5-36) можно заменить более простым

$$\sigma = \sigma_0 e^{-\frac{\Delta E_g}{2k_0 T}}. \quad (5-38)$$

В примесном полупроводнике, как следует из рис. 5-3, при достаточно высоких температурах проводимость полупроводника является собственной, а при низких температурах — примесной. В области истощения примеси концентрация основных носителей остается постоянной, и проводимость меняется вследствие изменения подвижности с температурой.

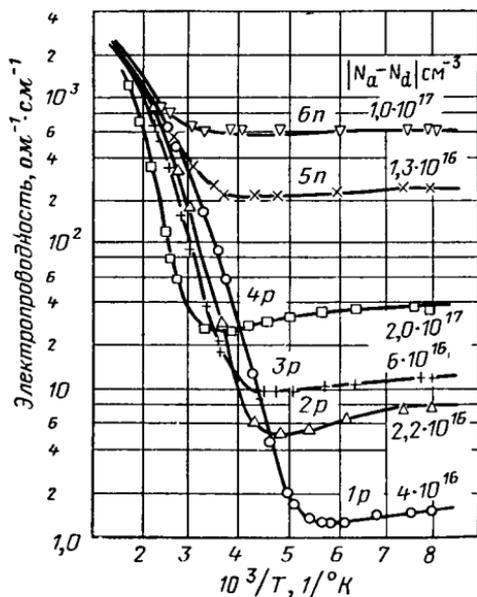


Рис. 5-3. Зависимость электропроводности от обратной температуры для двух n -типа и четырех p -типа образцов InSb .

описание его можно получить на основе решения кинетического уравнения Больцмана. Рассмотрение эффекта Холла, когда на носители заряда действует сила

$$\mathbf{F} = e \{ \mathcal{E} + [\mathbf{v}\mathbf{B}] \}, \quad (5-39)$$

проведем для слабого магнитного поля. Под слабым магнитным полем понимают такое поле, для которого время релаксации τ носителя заряда много меньше его периода обращения T_c по круговой орбите в магнитном поле, т. е.

$$\tau \ll T_c. \quad (5-40)$$

Как известно, частота обращения носителя заряда с эффективной массой m^* по круговой орбите в магнитном поле (частота ω_c

5-3. ЭФФЕКТ ХОЛЛА

Кинетические эффекты, имеющие место при одновременном воздействии на вещество электрического \mathcal{E} и магнитного \mathbf{H} полей, называют гальваномагнитными эффектами. Эффект Холла является одним из гальваномагнитных эффектов. При рассмотрении явления Холла в гл. 1 (§ 1-4) не учитывалось явление рассеяния. Более строгое

циклотронного резонанса) равна:

$$\omega_c = \frac{2\pi}{T_c} = \frac{e}{m^*} B. \quad (5-41)$$

Используя формулу (5-41), условие слабого магнитного поля (5-40) можно записать в виде

$$\tau \ll \frac{m^*}{e} \frac{2\pi}{B}. \quad (5-42)$$

Принимая во внимание, что $e\tau/m^*$ представляет собой подвижность носителей заряда μ , и усиливая неравенство (5-42), получаем:

$$\mu B \ll 1. \quad (5-43)$$

Это условие обычно и служит критерием слабого магнитного поля.

Рассмотрение эффекта Холла проведем для однородного полупроводника со сферическими изоэнергетическими поверхностями, находящегося в постоянных электрическом и магнитном полях.

Чтобы найти постоянную Холла R вычислим плотность тока

$$\mathbf{j} = \mathbf{j}_n + \mathbf{j}_p = -\frac{e\hbar}{4\pi^3 m_n^*} \int_{(V_B)} \kappa f_1^{(n)}(\kappa) d\tau_\kappa + \frac{e\hbar}{4\pi^3 m_p^*} \int_{(V_i)} \kappa' f_1^{(p)}(\kappa') d\tau_{\kappa'}. \quad (5-44)$$

Для нахождения $f_1(\kappa)$ используем (4-33), которое с учетом (5-39) для однородного полупроводника запишем:

$$\frac{e}{\hbar} \{ \mathcal{E} + [\mathbf{vB}] \} \nabla_\kappa f = -\frac{f_1(\kappa)}{\tau}. \quad (5-45)$$

Представим $f_1(\kappa)$ в виде

$$f_1(\kappa) = -\frac{\partial f_0}{\partial E} \kappa \chi(E), \quad (5-46)$$

где $\chi(E)$ — некоторый вектор, зависящий лишь от энергии носителя заряда.

Будем считать, что неравновесная функция $f(\kappa) = f_0 + f_1(\kappa)$ незначительно отличается от равновесной f_0 (слабые поля), и в уравнении (5-45) вместо f подставим f_0 везде, кроме члена, содержащего магнитное поле. Так как $[\mathbf{vB}] \mathbf{v} = 0$, то уравнение Больцмана будет:

$$\frac{e}{\hbar} \mathcal{E} \nabla_\kappa f_0 + \frac{e}{\hbar} [\mathbf{vB}] \nabla_\kappa f_1(\kappa) = -\frac{f_1(\kappa)}{\tau}. \quad (5-47)$$

Здесь

$$\nabla_\kappa f_0 = \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial \kappa} = \frac{\partial f_0}{\partial E} \cdot \hbar \mathbf{v};$$

$$\nabla_\kappa f_1(\kappa) = -\nabla_\kappa \left\{ \frac{\partial f_0}{\partial E} \kappa \chi \right\} = -\frac{\partial f_0}{\partial E} \chi - \left\{ \hbar \kappa \frac{\partial}{\partial E} \left(\frac{\partial f_0}{\partial E} \chi \right) \right\} \mathbf{v}. \quad (5-48)$$

Как известно, $[\mathbf{vB}] \chi = [\mathbf{B}\chi] \mathbf{v}$, и, если учесть выражения (5-46) и (5-48), уравнение (5-47) примет вид:

$$f_1(\kappa) = -\frac{\partial f_0}{\partial E} \kappa \chi = e\tau \frac{\partial f_0}{\partial E} \left\{ \mathcal{E} - \frac{1}{\hbar} [\mathbf{B}\chi] \right\} \mathbf{v}. \quad (5-49)$$

Следовательно,

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathcal{E} - \frac{1}{\hbar} [\mathbf{B}\chi_n] \right\}; \quad (5-50)$$

$$\chi_p = \frac{e\hbar\tau_h}{m_p^*} \left\{ \mathcal{E} - \frac{1}{\hbar} [\mathbf{B}\chi_p] \right\}. \quad (5-51)$$

Решая векторные уравнения (5-50) и (5-51) методом последовательных приближений, получаем для случая слабого магнитного поля

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B}\mathcal{E}] \right\}; \quad (5-52)$$

$$\chi_p = \frac{e\hbar\tau_h}{m_p^*} \left\{ \mathcal{E} - \frac{e\tau_h}{m_p^*} [\mathbf{B}\mathcal{E}] \right\}. \quad (5-53)$$

Находим теперь электронную составляющую тока, используя выражения (5-46) и (5-52):

$$\mathbf{j}_n = -\frac{e^2\hbar^2}{4\pi^3 m_n^{*3}} \int_{(V_B)} \frac{\partial f_0}{\partial E} \left\{ \tau_e(\kappa\mathcal{E}) + \frac{e\tau_e^2}{m_n^*} \kappa [\mathbf{B}\mathcal{E}] \right\} \kappa d\tau_\kappa. \quad (5-54)$$

Если учесть равенства (5-9), (5-11) и (5-13), то уравнение (5-54) запишется в виде

$$\mathbf{j}_n = \frac{e^2 n}{m_n^*} \mathcal{E} \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha + \frac{e^3 n}{m_n^{*2}} [\mathbf{B}\mathcal{E}] \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-\alpha} \alpha^{3/2} d\alpha. \quad (5-55)$$

Вводя среднее значение времени релаксации для электронов

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha \quad (5-56)$$

и среднее значение квадрата времени релаксации

$$\langle \tau_e^2 \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-\alpha} \alpha^{3/2} d\alpha \quad (5-57)$$

и учитывая, что $\mu_n = \frac{e \langle \tau_e \rangle}{m_n^*}$, получаем для плотности электронной составляющей тока

$$\mathbf{j}_n = en\mu_n \mathcal{E} + \frac{\langle \tau_e^2 \rangle}{|\langle \tau_e \rangle|^2} en\mu_n^2 [\mathbf{B}\mathcal{E}] = en\mu_n \mathcal{E} + r_n en\mu_n^2 [\mathbf{B}\mathcal{E}], \quad (5-58)$$

где

$$r_n = \frac{\langle \tau_e^2 \rangle}{|\langle \tau_e \rangle|^2}. \quad (5-59)$$

По аналогии плотность дырочной составляющей тока запишется в виде

$$\mathbf{j}_p = ep\mu_p \mathcal{E} - r_p ep\mu_p^2 [\mathbf{B}\mathcal{E}], \quad (5-60)$$

где

$$r_p = \frac{\langle \tau_h^2 \rangle}{|\langle \tau_h \rangle|^2}. \quad (5-61)$$

Полная плотность тока

$$\mathbf{j} = e(\rho\mu_p + n\mu_n)\mathcal{E} - r_n e \left(\frac{r_p}{r_n} \rho\mu_p^2 - n\mu_n^2 \right) [\mathbf{B}\mathcal{E}]. \quad (5-62)$$

В уравнении (5-62) коэффициенты r_n и r_p зависят от механизма рассеяния носителей заряда, который определяет время релаксации τ (к). В случае упругого рассеяния, когда $\kappa' = \kappa$ и $E' = E$, для любого механизма рассеяния отношение

$$\frac{r_p}{r_n} = \frac{\langle \tau_h^2 \rangle_n |\langle \tau_e \rangle|^2}{|\langle \tau_h \rangle|^2 \langle \tau_e^2 \rangle} = 1. \quad (5-63)$$

Обозначим $r_n = r$ и, учитывая равенство (5-63), будем иметь:

$$\mathbf{j} = e(n\mu_n + \rho\mu_p)\mathcal{E} - re(\rho\mu_p^2 - n\mu_n^2)[\mathbf{B}\mathcal{E}]. \quad (5-64)$$

Пусть электрический ток направлен 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 $B_x = B_y = 0$; при этом равенство (5-64) распадается на два

$$j_x = e(\rho\mu_p + n\mu_n)\mathcal{E}_x + re(\rho\mu_p^2 - n\mu_n^2)B\mathcal{E}_y = j, \quad (5-65)$$

$$j_y = e(\rho\mu_p + n\mu_n)\mathcal{E}_y - re(\rho\mu_p^2 - n\mu_n^2)B\mathcal{E}_x = 0. \quad (5-66)$$

Решая уравнения (5-65) и (5-66) относительно \mathcal{E}_y , в случае слабого магнитного поля получаем для напряженности холловского поля

$$\mathcal{E}_y = \frac{r}{e} \frac{\rho\mu_p^2 - n\mu_n^2}{(\rho\mu_p + n\mu_n)^2} jB = RjB. \quad (5-67)$$

Отсюда найдем коэффициент Холла

$$R = \frac{r}{e} \frac{\rho\mu_p^2 - n\mu_n^2}{(\rho\mu_p + n\mu_n)^2}. \quad (5-68)$$

Таким образом, коэффициент Холла зависит от механизма рассеяния. Это учитывается множителем r .

Определим величину r разных механизмов рассеяния. При рассеянии на акустических колебаниях решетки время релаксации с учетом выражения (4-162) и $E = k_0 T \alpha$ будет равно:

$$\tau = \frac{\tau_0}{m_n^{*3/2}} T^{-1} E^{-1/2} = \frac{\tau_0}{m^{*3/2} k_0^{1/2} T^{3/2}} \alpha^{-1/2}, \quad (5-69)$$

и на основании формулы (5-59) найдем:

$$r = \frac{3\pi}{8}. \quad (5-70)$$

При рассеянии на ионах примеси согласно выражению (4-68)

$$r = \frac{315\pi}{512} = 1,93. \quad (5-71)$$

Поскольку рассеяние зависит от температуры, то при определении коэффициента Холла при низких температурах, когда основную роль в рассеянии играют ионы примеси, необходимо полагать

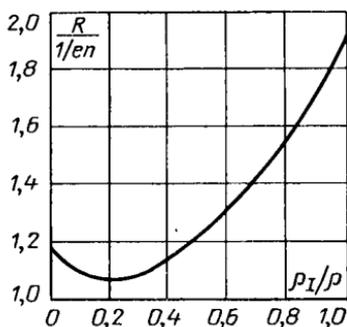


Рис. 5-4. Зависимость величины коэффициента $r = R / \frac{1}{en}$ от отношения ρ_I / ρ .

$r = 1,93$. Для температур, при которых имеет место рассеяние на акустических фонах, $r = 3\pi/8$. В том случае, когда в рассеянии одновременно участвуют как колебания решетки, так и ионы примеси, r выражается сложной функцией. На рис. 5-4 показана зависимость величины r от отношения сопротивления, обусловленного рассеянием на примеси, к полному сопротивлению.

Для электронного полупроводника коэффициент Холла на основании формулы (5-68) имеет вид:

$$R = -r \frac{1}{en}. \quad (5-72)$$

В то же время проводимость такого полупроводника

$$\sigma = en\mu_d, \quad (5-73)$$

где

$$\mu_d = \frac{\sigma}{en} = \frac{e \langle \tau \rangle}{m^*} \quad (5-74)$$

есть дрейфовая или омическая подвижность. Следовательно,

$$\frac{1}{r} |R| \sigma = \mu_d. \quad (5-75)$$

Но с помощью равенства

$$|R| \sigma = \mu_H \quad (5-76)$$

можно определить величину μ_H , имеющую размерность подвижности, а поэтому называемую холловской подвижностью. Из сравнения равенств (5-75) и (5-76) следует, что

$$\mu_H = r\mu_d, \quad (5-77)$$

т. е. холловская подвижность, определяющая угол Холла, пропорциональна дрейфовой подвижности.

В случае, когда время релаксации $\tau = \text{const}$, т. е. не зависит от энергии,

$$r = 1, \quad (5-78)$$

поэтому

$$\mu_H = \mu_d. \quad (5-79)$$

Это имеет место у металлов и вырожденных полупроводников.

5-4. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ

К термоэлектрическим явлениям относятся три эффекта: эффект Зеебека — возникновение термоэлектрического напряжения в цепи, состоящей из двух проводников, места соединения которых находятся при различных температурах; эффект Пельтье — выделение или поглощение тепла на контакте двух проводников при прохождении электрического тока и эффект Томсона — выделение или поглощение тепла в объеме проводника при прохождении электрического тока и наличии градиента температуры.

Рассмотрим однородный полупроводник при наличии градиента температуры. Если полупроводник нагрет неравномерно, то средняя энергия носителей заряда и их концентрация в нем будут больше там, где выше температура. Следовательно, градиент температуры в однородном полупроводнике приводит к градиенту средней энергии носителей заряда и градиенту их концентрации, вследствие чего возникнет диффузионный поток носителей заряда, т. е. появится ток. Но в разомкнутой цепи в стационарном состоянии плотность тока во всех точках образца равна нулю. В данном случае это возможно благодаря появлению в результате разделения заряда электрического поля, которое вызовет в каждой точке полупроводника ток, компенсирующий поток носителей, обусловленный градиентом температуры. На образце возникнет электродвижущая сила, которую называют термоэлектродвижущей силой (термо-э. д. с.).

Термоэлектродвижущая сила, отнесенная к единичной разности температуры, называется дифференциальной термо-э. д. с. и обозначается α .

Определение термо-э. д. с. проведем с помощью кинетического уравнения (4-33) для невырожденного полупроводника, имеющего сферические изоэнергетические поверхности для электронов $E = \hbar^2 \kappa^2 / 2m_n^*$ и дырок $E' = \hbar^2 \kappa'^2 / 2m_p^*$. Если химический потенциал для электронов равен F , то для дырок он будет $F' = F - \Delta E_g$, где ΔE_g — ширина запрещенной зоны. Поскольку полупроводник невырожденный, то равновесная функция распределения для электронов $f_0 = e^{-\frac{E-F}{k_0T}}$ и для дырок $f_{0p} = e^{-\frac{F'-E'}{k_0T}} = e^{-\frac{E'+F+\Delta E_g}{k_0T}}$.

Поскольку $n = N_c e^{\frac{E-F}{k_0T}}$, а $p = N_v e^{\frac{F'-E'}{k_0T}}$, находим:

$$\frac{\partial f_0}{\partial E} = -\frac{1}{k_0T} e^{-\frac{E-F}{k_0T}} = -\frac{n}{k_0TN_c} e^{-\frac{E}{k_0T}}; \quad (5-80)$$

$$\frac{\partial f_0}{\partial T} = \frac{E-F}{k_0T^2} e^{-\frac{E-F}{k_0T}} = \frac{F-E}{T} \frac{\partial f_0}{\partial E}; \quad (5-81)$$

$$\frac{\partial f_0}{\partial F} = \frac{1}{k_0T} e^{-\frac{E-F}{k_0T}} = -\frac{\partial f_0}{\partial E}. \quad (5-82)$$

При малой величине отклонения $f_1(\mathbf{k}) = f(\mathbf{k}) - f_0$ в левой части уравнения (4-33) можно заменить f на f_0 . Тогда, учитывая уравнения (5-81) и (5-82), получаем:

$$\nabla_r f \approx \nabla_r f_0 = \frac{\partial f_0}{\partial T} \nabla T + \frac{\partial f_0}{\partial F} \nabla F = \frac{\partial f_0}{\partial E} \left(\frac{F-E}{T} \nabla T - \nabla F \right); \quad (5-83)$$

$$\nabla_{\mathbf{k}} f \approx \nabla_{\mathbf{k}} f_0 = \frac{\partial f_0}{\partial E} \nabla_{\mathbf{k}} E = \hbar \frac{\partial f_0}{\partial E} \mathbf{v}_n. \quad (5-84)$$

Подставляя выражения (5-83) и (5-84) в кинетическое уравнение (4-33) и ограничиваясь случаем, когда на электрон действует только электрическое поле $\mathcal{E} = -\nabla\varphi$ (φ — электростатический потенциал), и учитывая, что $F = -e\mathcal{E} = e\nabla\varphi$, будем иметь:

$$\mathbf{v}_n \frac{\partial f_0}{\partial E} \left\{ \frac{F-E}{T} \nabla T - \nabla F \right\} + e \frac{\partial f_0}{\partial E} \mathbf{v}_n \nabla \varphi = - \frac{f_1^{(n)}(\mathbf{k})}{\tau_e(\mathbf{k})}. \quad (5-85)$$

Из этого уравнения можно определить:

$$f_1^{(n)}(\mathbf{k}) = - \tau_e \frac{\partial f_0}{\partial E} \left\{ \frac{F-E}{T} \nabla T - \nabla(F - e\varphi) \right\} \mathbf{v}_n. \quad (5-86)$$

Проведя подобные выкладки для дырок, найдем:

$$f_1^{(p)}(\mathbf{k}') = \tau_h \frac{\partial f_0}{\partial E'} \left\{ \frac{E' + F + \Delta E_g}{T} \nabla T - \nabla(F - e\varphi) \right\} \mathbf{v}_p. \quad (5-87)$$

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 (4-163) и время релаксации можно выразить через l , т. е.

$$\tau = \frac{l}{v} = \frac{m_n^* l}{\hbar} \kappa^{-1}. \quad (5-88)$$

Электронную составляющую тока с учетом (5-86), (5-88) и (5-11) можно представить в виде

$$\begin{aligned} \mathbf{j}_n = & \frac{e\hbar l}{3\pi^2 m_n^*} \left\{ \frac{F}{T} \nabla T - \nabla(F - e\varphi) \right\} \int_0^\infty \frac{\partial f_0}{\partial E} \kappa^3 d\kappa - \\ & - \frac{e\hbar l}{3\pi^2 m_n^* T} \nabla T \int_0^\infty \frac{\partial f_0}{\partial E} E \kappa^3 d\kappa. \end{aligned} \quad (5-89)$$

Если ввести $E = k_0 T \alpha$, то $\kappa^3 d\kappa = \frac{(2m_n^* k_0 T)^2}{2\hbar^4} \alpha d\alpha$, при этом, учитывая формулы (5-80) и (5-88), получаем:

$$\mathbf{j}_n = - \frac{4enl}{3(2\pi m_n^* k_0 T)^{1/2}} \left\{ \frac{F}{T} \nabla T - \nabla(F - e\varphi) - 2k_0 \nabla T \right\}. \quad (5-90)$$

Используя соотношение подвижности носителей заряда (5-26), выраженную через длину свободного пробега, формулу (5-90) можно записать в виде

$$j = n\mu_n \left\{ \nabla (F - e\varphi) + \left(2k_0 - \frac{F}{T} \right) \nabla T \right\}. \quad (5-91)$$

Проведя расчет для дырочной составляющей тока, найдем:

$$j_p = p\mu_p \left\{ \nabla (F - e\varphi) - \left(2k_0 + \frac{F + \Delta E_g}{T} \right) \nabla T \right\}. \quad (5-92)$$

Следовательно, полная плотность тока будет равна:

$$j = n\mu_n \left\{ \nabla (F - e\varphi) + \left(2k_0 - \frac{F}{T} \right) \nabla T \right\} + p\mu_p \left\{ \nabla (F - e\varphi) - \left(2k_0 + \frac{F + \Delta E_g}{T} \right) \nabla T \right\}. \quad (5-93)$$

Из соотношения (5-93) следует, что в полупроводнике при наличии градиента температуры возникает электронный и дырочный токи, которые обусловлены действием градиента электрохимического потенциала $\nabla (F - e\varphi)$ и градиента температуры ∇T . Для нахождения термо-э. д. с. 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 равенства (5-93), приравняв его нулю. Но при измерении термо-э. д. с. на границах полупроводника и измерительных металлических электродов существуют контактные разности потенциалов (см. § 8-4), которые не равны друг другу вследствие существующего температурного градиента. Измерительный прибор отметит э. д. с., равную термо-э. д. с. полупроводника и разности контактных потенциалов измерительных электродов. Чтобы исключить последнее слагаемое, следует определить термо-э. д. с. как градиент электрохимического потенциала $\nabla (F - e\varphi)$. Величина $\left| \nabla \frac{F}{e} \right|$ будет равна разности контактных потенциалов в граничных точках, если считать, что полупроводник и металл на контакте находятся в термодинамическом равновесии, тогда дифференциальная термо-э. д. с. α определяется как

$$\alpha = \frac{\left| \nabla \left(\varphi - \frac{F}{e} \right) \right|}{|\nabla T|}. \quad (5-94)$$

Полагая ток в соотношении (5-93) равным нулю, получаем для дифференциальной термо-э. д. с. полупроводника

$$\alpha = \frac{k_0}{e} \frac{1}{n\mu_n + p\mu_p} \left\{ n\mu_n \left(2 - \frac{F}{k_0 T} \right) - p\mu_p \left(2 + \frac{F + \Delta E_g}{k_0 T} \right) \right\}. \quad (5-95)$$

Если воспользоваться формулами для концентрации электронов и дырок невырожденного полупроводника, то выражение (5-95)

можно записать в виде

$$\alpha = \frac{k_0}{e} \frac{1}{n\mu_n + p\mu_p} \left\{ n\mu_n \left[2 + \ln \frac{2(2\pi m_n^* k_0 T)^{3/2}}{h^3 n} \right] - \right. \\ \left. - p\mu_p \left[2 + \ln \frac{2(2\pi m_p^* k_0 T)^{3/2}}{h^3 p} \right] \right\}. \quad (5-96)$$

Таким образом, термо-э. д. с. полупроводника определяется двумя слагаемыми, каждое из которых соответствует вкладу, вносимому электронами и дырками, причем эти слагаемые имеют противоположные знаки.

В случае электронного полупроводника на горячем торце образца возникает положительный объемный заряд, поскольку электроны диффундируют от горячего торца к холодному. В дырочном полупроводнике, наоборот, на горячем торце возникает отрицательный объемный заряд. Таким образом, *если полупроводник примесной, то направление внутреннего электрического поля и полярность термо-э. д. с. определяются знаком носителей заряда, и, следовательно, по знаку термо-э. д. с. можно определить тип примесной проводимости исследуемого образца.* В случае смешанной проводимости знак термо-э. д. с. определяется не только соотношением концентраций носителей, но и соотношением их подвижностей. В свою очередь величина термо-э. д. с. примесных полупроводников уменьшается с ростом содержания примеси. При нагревании примесного полупроводника его термо-э. д. с. также уменьшается.

Для собственного полупроводника $n = p = n_i$, $F = -\frac{\Delta E_g}{2}$ и формулу (5-95) можно переписать в виде

$$\alpha = \frac{k_0}{e} \frac{b-1}{b+1} \left(2 + \frac{\Delta E_g}{2k_0 T} \right), \quad (5-97)$$

где $b = \mu_n / \mu_p$.

Величина термо-э. д. с. собственного полупроводника согласно выражению (5-97) определяется лишь шириной запрещенной зоны и соотношением подвижности электронов и дырок.

Изменение механизма рассеяния в полупроводнике, которое может иметь место при различных температурах, влияет на термо-э. д. с. через подвижность и сказывается на величине первого слагаемого (в квадратных скобках) выражения (5-96). Например, его величина для рассеяния на акустических колебаниях и ионах примеси в атомных кристаллах, оптических колебаниях в ионных кристаллах выше и ниже температуры Дебая принимает соответственно значения 2, 4, 3 и 5/2.

На основании выражения (5-96) нельзя объяснить большое возрастание термо-э. д. с., которое обнаруживается иногда в чистом веществе при очень низких температурах. Этот эффект определяется особым видом электронно-фононного взаимодействия, которое получило название увлечения носителей заряда ф о н о н а м и.

При наличии градиента температуры в полупроводнике имеет место анизотропия в распространении фононов, которые преимущественно движутся от горячего к холодному краю образца. При каждом столкновении электрона с фононом, которое сопровождается уничтожением фонона, его импульс передается электрону. Поскольку фононов с импульсом, направленным от горячего к холодному краю образца, больше, чем противоположно направленным импульсом, то часть импульса будет передаваться электронам. Вследствие чего 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 уравнивает силу, действующую на электроны со стороны

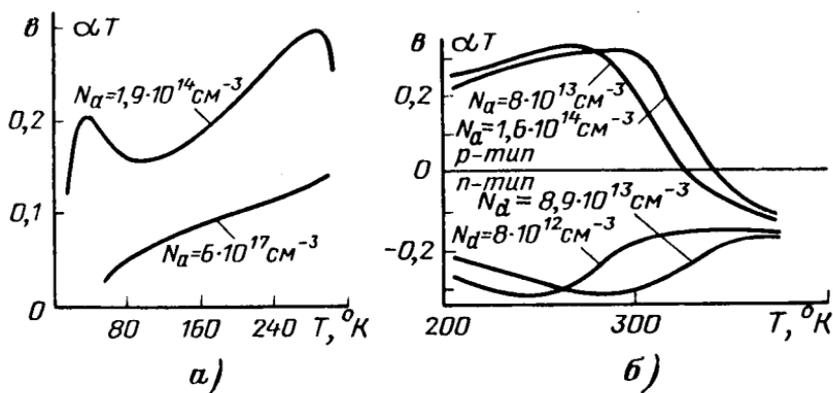


Рис. 5-5. Термо-э. д. с. германия при низких (а) и высоких (б) температурах.

фононного потока. Эта разность потенциалов и представляет собой дополнительную термо-э. д. с. Эффект увлечения приводит к значительному росту термо-э. д. с. 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-э. д. с. полупроводников хорошо согласуются с экспериментальными результатами, полученными для германия как при высоких, так и при низких температурах (рис. 5-5).

У металлов концентрация электронов не зависит от температуры, а их средняя энергия¹ равна:

$$\langle E \rangle = \frac{5}{3} F_0 \left[1 + \frac{5\pi^2}{12} \left(\frac{k_0 T}{F_0} \right)^2 \right], \quad (5-98)$$

¹ Зейтц Ф., Современная теория твердого тела, Гостехиздат, 1949, гл. 4, § 26.

где F_0 — уровень Ферми при температуре абсолютного нуля. Так как для металлов $F_0 \approx 5$ эв, то второй член в выражении (5-98) мал и энергия электронов в металле практически не зависит от температуры. В силу этого термо-э. д. с. металлов мала по сравнению с полупроводниками. Для металлов $\alpha \approx 1 \div 10$ мкв/град, а у полупроводников она достигает величины порядка 100—1 000 мкв/град. Такое различие в величинах термо-э. д. с. приводит к тому, что при наличии в цепи полупроводника металла полная термо-э. д. с. слабо зависит от свойств металла, а поэтому, как правило, приводят значение термо-э. д. с. для полупроводника, не указывая, относительно какого металла она определена.

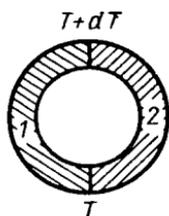


Рис. 5-6. Замкнутая цепь двух проводников.

В проводниках наряду с термо-э. д. с. существуют еще два термоэлектрических эффекта — эффект Пельтье и эффект Томсона.

Явление Пельтье заключается в том, что при прохождении электрического тока через контакт двух веществ в нем, кроме тепла Джоуля, в зависимости от направления тока выделяется или поглощается тепло. Количество тепла при эффекте Пельтье пропорционально плотности тока и времени

$$Q_P = Pjt, \quad (5-99)$$

где P — коэффициент Пельтье.

Если вдоль однородного проводника существует градиент температуры, то при прохождении тока в объеме материала выделяется (или поглощается) тепло Томсона, количество которого пропорционально количеству протекающего электричества и перепаду температур

$$Q_T = It \int_T^{T_1} \tau_T dT. \quad (5-100)$$

Коэффициент τ_T называется коэффициентом Томсона.

Коэффициенты α , P и τ_T связаны между собой соотношениями Томсона. Для их вывода проведем элементарное рассмотрение термоэлектрических эффектов. Для этого составим замкнутую цепь из двух проводников, как это изображено на рис. 5-6. Пусть между контактами 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 тепло Пельтье в количестве $P'dI$, а на другом поглощается PdI . В результате в цепи выделяется тепло Пельтье, и если P' мало отличается от P , то количество этого тепла будет:

$$dQ_{II} = (P' - P) dI = \frac{\partial P}{\partial T} dI dT. \quad (5-101)$$

Естественно, наличие электрического тока и градиента температуры вызовет появление тепла за счет эффекта Томсона, равного за 1 сек

$$dQ_T = (\tau_{T_1} - \tau_{T_2}) dl dT. \quad (5-102)$$

Эти термодинамические явления можно считать обратимыми, так как они зависят от направления тока и градиента температуры. Но в данной замкнутой цепи имеют место и необратимые процессы — теплопроводность и выделение тепла Джоуля. В стационарных условиях теплопроводность не изменяет общего количества тепла в системе, имеет место только перенос тепла из одного участка в другой. Тепло Джоуля пропорционально квадрату тока dI^2 , поэтому при малых значениях dI и dT эти необратимые эффекты можно не учитывать. На основании закона сохранения энергии можно написать:

$$\alpha dl dT = \frac{\partial \Pi}{\partial T} dl dT + (\tau_{T_1} - \tau_{T_2}) dl dT, \quad (5-103)$$

откуда найдем:

$$\alpha = \frac{\partial \Pi}{\partial T} + (\tau_{T_1} - \tau_{T_2}). \quad (5-104)$$

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 замкнутой системы при обратимых процессах равно нулю, имеем:

$$\frac{\Pi' dl}{T'} - \frac{\Pi dl}{T} + dl \int_T^{T'} \frac{\tau_{T_1} - \tau_{T_2}}{T} dT = 0. \quad (5-105)$$

Если за начало отсчета энтропии взять ее значение на горячем конце ($\Pi' dl/T'$), то, продифференцировав (5-105) по T , получим:

$$\frac{d}{dT} \left(\frac{\Pi}{T} \right) + \frac{\tau_{T_1} - \tau_{T_2}}{T} = 0, \quad (5-106)$$

откуда найдем:

$$\frac{d\Pi}{dT} + (\tau_{T_1} - \tau_{T_2}) = \frac{\Pi}{T}. \quad (5-107)$$

Сравнивая уравнения (5-107) и (5-104), видим, что

$$\Pi = \alpha T. \quad (5-108)$$

Если продифференцируем равенство (5-108) по T , то найдем:

$$\frac{d\alpha}{dT} = \frac{1}{T} \frac{d\Pi}{dT} - \frac{\Pi}{T^2} = \frac{1}{T} \left(\frac{d\Pi}{dT} - \frac{\Pi}{T} \right). \quad (5-109)$$

Учитывая уравнение (5-107), получаем:

$$\tau_{T_2} - \tau_{T_1} = \tau_T = T \frac{d\alpha}{dT}. \quad (5-110)$$

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.

5-5. ЯВЛЕНИЯ ПЕРЕНОСА В СИЛЬНЫХ ЭЛЕКТРИЧЕСКИХ ПОЛЯХ

При выводе кинетического уравнения мы предполагали, что время релаксации τ не зависит от величины поля. Используя кинетическое уравнение для невырожденного полупроводника со сферическими изоэнергетическими поверхностями, показали, что плотность тока подчиняется закону Ома, т. е. пропорциональна напряженности электрического поля

$$\mathbf{j} = en\mathbf{v}_d = en\mu_0\mathcal{E} = \sigma_0\mathcal{E}, \quad (5-111)$$

где μ_0 , σ_0 — подвижность и проводимость в слабых полях.

В электрическом поле с напряженностью \mathcal{E} электрон на длине свободного пробега l приобретает энергию, равную $e\mathcal{E}l$. При столкновении с акустическим фононом электрон эту энергию передает решетке, приходя в тепловое равновесие с ней. Но при столкновении с фононом электрон может либо передать энергию фонону, либо получить энергию от фонона. При этом изменение энергии электрона составляет величину ΔE . Если из общего числа столкновений подсчитать число столкновений, идущих с передачей энергии фонону, то на одно такое столкновение приходится $(2N_q + 1)$ столкновений, поэтому относительная потеря энергии при одном столкновении согласно выражению (4-165) будет равна: $\delta = \frac{\Delta E}{k_0 T} \frac{1}{k_0 T} \hbar\omega_q \times \times \frac{1}{2N_q + 1} = \frac{1}{k_0 T} \hbar\omega_q \operatorname{th} \frac{\hbar\omega_q}{2k_0 T}$.

Следовательно, изменение энергии электрона при столкновении равно $\Delta E = \delta k_0 T$. Соударение будет упругим в том случае, если

$$\frac{e\mathcal{E}l}{\Delta E} \ll 1. \quad (5-112)$$

Это значит, что закон Ома выполняется в том случае, если

$$e\mathcal{E}l \ll k_0 T \delta \quad (5-113)$$

или для величины поля

$$\mathcal{E} \ll \delta \frac{k_0 T}{el}. \quad (5-114)$$

При этом условии распределение скоростей носителей заряда остается почти максвелловским с температурой, равной температуре решетки. В этом случае на хаотическое движение носителей накладывается медленный их дрейф, обусловленный действием слабого внешнего электрического поля. При этом скорость дрейфа значи-

тельно меньше тепловой скорости

$$v_d \ll v_0. \quad (5-115)$$

В сильных электрических полях имеет место нарушение закона Ома. Это может происходить по двум причинам. Во-первых, в сильных электрических полях скорость дрейфа носителей становится соизмеримой с тепловой скоростью. Это скажется на процессе рассеяния, что проявится 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 может обуславливаться рядом причин. Такими эффектами являются ударная ионизация, эффект Зинера и электростатическая ионизация. Рассмотрим каждое из этих явлений в отдельности.

Электропроводность в сильном электрическом поле. В сильных электрических полях, для которых не выполняется условие (5-114), носители заряда при столкновении с акустическими фононами не успевают отдавать набираемую ими в поле энергию, и их средняя энергия повышается и электронно-дырочный газ «нагревается». Рост энергии носителей будет продолжаться до тех пор, пока не установится новое стационарное состояние, при котором энергия, приобретаемая носителем в поле за единицу времени $e\mathcal{E}v_d$, сравняется с энергией $\frac{\Delta E}{\tau} = \frac{\delta k_0 T}{\tau}$, теряемой этим носителем в результате столкновения, т. е.

$$e\mathcal{E}v_d = \delta \frac{k_0 T}{\tau}. \quad (5-116)$$

При рассеянии на тепловых колебаниях l не зависит от энергии, и время свободного пробега $\tau = l/v_0$. Следовательно, дрейфовая скорость носителей заряда согласно формуле (5-116) равна:

$$v_d = \frac{\delta k_0 T}{e\mathcal{E}\tau} = \frac{\delta k_0 T v_0}{e\mathcal{E}l}. \quad (5-117)$$

В то же время дрейфовая скорость

$$v_d = \frac{e\tau}{m^*} \mathcal{E} = \frac{el}{m^* v_0} \mathcal{E}. \quad (5-118)$$

Из сравнения формул (5-117) и (5-118) следует, что

$$(e\mathcal{E}l)^2 = 2\delta k_0 T \frac{m^* v_0^2}{2}. \quad (5-119)$$

Но так как $m^* v_0^2/2 = k_0 T$, то из соотношения (5-119) найдем напряженность поля

$$\mathcal{E} = \frac{k_0 T}{el} (2\delta)^{1/2} = \mathcal{E}_{кр}. \quad (5-120)$$

при котором начинается «разогрев» электронно-дырочного газа. Это имеет место при условии, если дополнительная скорость, приобретаемая носителем заряда в поле, становится сравнимой с тепловой скоростью. Как следует из выражения (5-120), при уменьшении температуры величина критического поля уменьшается.

Из формулы (5-119) можно определить энергию носителей, выраженную через тепловую скорость

$$\left(\frac{m^* v_0^2}{2}\right)^2 = \frac{(e\mathcal{E}l)^2}{2\delta}. \quad (5-121)$$

Из равенства (5-121) найдем v_0 и, подставляя ее в формулу (5-117), получим:

$$v_d = \frac{e\mathcal{E}l}{m^*} \left[\frac{m^* (2\delta)^{1/2}}{2e\mathcal{E}l} \right]^{1/2} = \left[\frac{el}{m^*} \left(\frac{\delta}{2} \right)^{1/2} \right]^{1/2} \mathcal{E}^{1/2}. \quad (5-122)$$

Таким образом, в сильных электрических полях, когда величина поля больше критического, т. е. когда

$$\mathcal{E} > \mathcal{E}_{кр} \quad \text{или} \quad \mathcal{E} > \frac{k_0 T}{el} (2\delta)^{1/2}, \quad (5-123)$$

скорость движения носителей заряда будет увеличиваться как

$$v_d \sim \mathcal{E}^{1/2}, \quad (5-124)$$

а время релаксации соответственно будет уменьшаться как

$$\tau = \frac{l}{v_d} \sim \frac{1}{\mathcal{E}^{1/2}}, \quad (5-125)$$

откуда следует, что

$$\mu = \mu_0 \mathcal{E}^{-1/2}; \quad (5-126)$$

$$\sigma = \sigma_0 \mathcal{E}^{-1/2} \quad \text{и} \quad j = \sigma_0 V \sqrt{\mathcal{E}}, \quad (5-127)$$

т. е. при электрических полях $\mathcal{E} > \mathcal{E}_{кр}$ закон Ома не выполняется. Например, в германии отклонения от омической проводимости при комнатной температуре наступают при напряженностях поля порядка 200 в/см.

В области еще более сильных электрических полей или при достаточно высоких температурах, когда $T > \theta_0$, энергия носителей заряда увеличивается настолько, что возможно рассеяние их на оптических колебаниях. В этом случае при столкновении с оптическим фоном носитель отдает энергию, равную $\hbar\omega_0$, и скорость потери энергии сильно возрастает. В результате устанавливается новое стационарное состояние, при котором энергия $e\mathcal{E}v_d$, набираемая носителем в поле за единицу времени, равна энергии $\Delta E/\tau$, теряемой этим носителем за то же время, т. е.

$$e\mathcal{E}v_d = \frac{\Delta E}{\tau}. \quad (5-128)$$

Но энергия, которую носитель теряет при одном столкновении, согласно выражению (4-165) равна:

$$\Delta E = \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2k_0T}, \quad (5-129)$$

поэтому

$$e\mathcal{E}v_d = \frac{v_0}{l} \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2k_0T}. \quad (5-130)$$

Если из формулы (5-118) найти величину \mathcal{E} и подставить ее в формулу (5-130), то получим дрейфовую скорость носителей заряда, равную

$$v_d = \left(\frac{\hbar\omega_0}{m^*} \operatorname{th} \frac{\hbar\omega_0}{2k_0T} \right)^{1/2}. \quad (5-131)$$

Предельное значение скорости дрейфа

$$v_{ds} = \left(\frac{\hbar\omega_0}{m^*} \right)^{1/2}. \quad (5-132)$$

Таким образом, в пределе плотность тока через образец стремится к насыщению. При этом ток насыщения равен:

$$j_s = env_{ds} = en \left(\frac{\hbar\omega_0}{m^*} \right)^{1/2}. \quad (5-133)$$

Это значит, что

$$\sigma = \sigma_0 \mathcal{E}^{-1}. \quad (5-134)$$

Из (5-130) и (5-131) следует, что при увеличении электрического поля до величины

$$\mathcal{E}'_{\text{кр}} = \frac{v_0}{el} \left(m^* \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2k_0T} \right)^{1/2} \quad (5-135)$$

в валентных кристаллах имеет место насыщение тока, при котором ток не зависит от поля. Оценки показывают, что насыщение наступает в полях порядка 10^4 в/см, а величина $v_{ds} = 10^7$ см/сек. Это согласуется с данными экспериментов. На рис. 5-7 представлена зависимость плотности тока от электрического поля для электронного германия при трех температурах.

В полярных кристаллах рассеяние на оптических колебаниях становится преобладающим уже при малых значениях энергии носителей заряда. При этом в полях, при которых энергия электрона $E \leq \hbar\omega_0$, подвижность, как и в валентных полупроводниках, уменьшается с ростом 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.

Эффект Ганна. В полупроводниках, зона проводимости которых имеет более одного минимума энергии, электрон с волновым век-

тором k , соответствующим одному из минимумов, при рассеянии может оказаться в состоянии с волновым вектором k' , принадле-

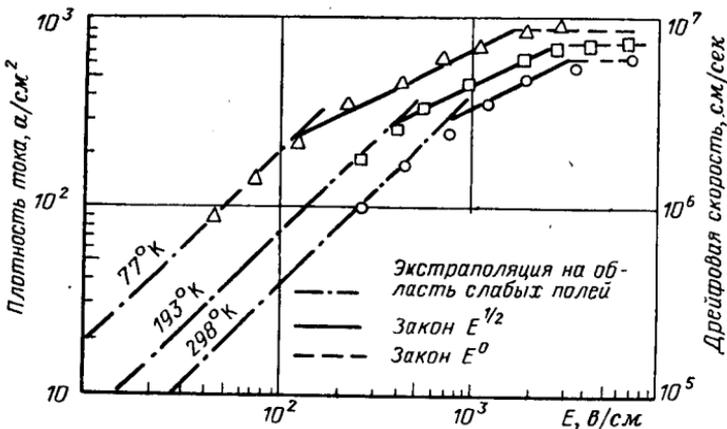


Рис. 5-7. Зависимость плотности тока от напряженности электрического поля в электронном германии.

жащим другому минимуму. В результате такого рассеяния будет иметь место переброс электронов из минимума в минимум. Такой вид рассеяния получил название «междолинного».

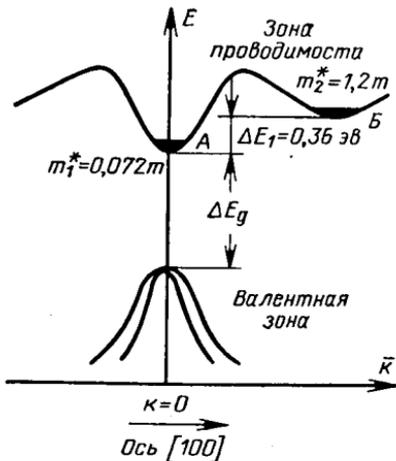


Рис. 5-8. Структура зон арсенида галлия в направлении оси [100].

электронов различны: легкие с эффективной массой $m_1^* = 0,072 m$ и тяжелые электроны с $m_2^* = 1,2 m$.

Легкие электроны характеризуются высокой подвижностью ($\mu_1 > 5000$ в/см²·сек), а тяжелые — низкой подвижностью ($\mu_2 \approx 100 \div 200$ в/см²·сек).

Междолинное рассеяние носителей заряда в определенных условиях может приводить к возникновению электрических колебаний в полупроводнике при приложении к нему постоянного напряжения достаточной величины. Это явление, названное эффектом Ганна, впервые наблюдалось на GaAs. На рис. 5-8 изображена зависимость энергии электрона в зоне проводимости GaAs от волнового вектора k для направления [100].

Существенным здесь является наличие двух минимумов A и B , разделенных зазором $\Delta E_1 = 0,36$ эв, в которых эффективные массы

При малых внешних электрических полях электроны будут находиться в термодинамическом равновесии с решеткой, имеющей температуру T_0 . Поскольку $k_0T \ll \Delta E_1$, то электроны в основном будут занимать энергетические уровни минимума A . Плотность тока будет определяться концентрацией n_0 и подвижностью μ_1 легких электронов:

$$j_n = en\mu_1 \mathcal{E}. \quad (5-136)$$

По мере возрастания напряженности электрического поля средняя энергия электронов повышается и становится возможным переход электронов во второй минимум B , где они будут обладать большой эффективной массой и малой подвижностью. Таким образом, с ростом напряженности поля будет иметь место изменение подвижности электронов, а следовательно, и их скорости. Это скажется на величине плотности тока. Скорость движения электрона, а значит, и плотность тока будут линейно возрастать с ростом напряженности поля до некоторого критического значения поля $\mathcal{E}_{кр}$.

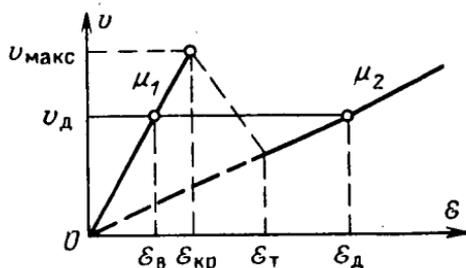


Рис. 5-9. Зависимость скорости движения носителей заряда в арсениде галлия от напряженности поля в образце.

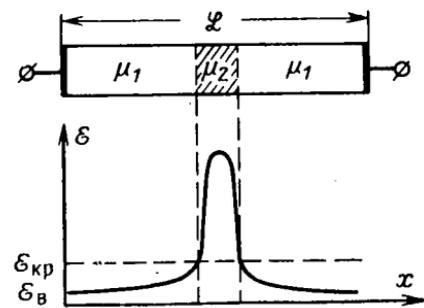


Рис. 5-10. Распределение электрического поля вдоль образца.

Скорости движения электронов и плотность тока будут линейно возрастать с ростом напряженности поля до некоторого критического значения поля $\mathcal{E}_{кр}$.

При $\mathcal{E} > \mathcal{E}_{кр}$ электроны, рассеиваясь, переходят в более высокий минимум B , где они будут обладать тяжелой массой m_2^* . Подвижность, а следовательно, и скорость движения электронов резко падают, и плотность тока резко понижается. На вольт-амперной характеристике появится падающий участок. Падающий участок будет иметь место и на графике зависимости

образца. Как только напряженность поля достигнет $\mathcal{E}_{кр}$, имеет место переход электронов в высокий минимум B и в области образца с повышенной напряженностью поля происходит возрастание концентрации тяжелых электронов. Подвижность электронов в этой части образца уменьшается, а сопротивление ее возрастает. Это приводит к возрастанию напряженности поля в этом месте образца, что в свою очередь вызывает более интенсивный переход электронов в верхний минимум, а справа и слева от этой области напряженность поля будет падать и распределение поля станет резко неоднородным (рис. 5-10). Эта зона, в которой сосредоточены тяжелые электроны, под действием электрического поля будет перемещаться вдоль образца с относительно низкой скоростью, обусловленной низкой подвижностью тяжелых электронов.

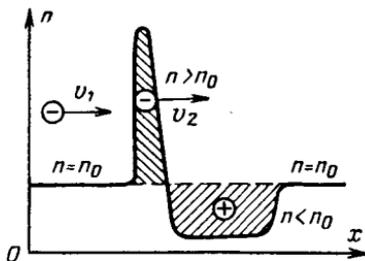


Рис. 5-11. Структура электрического домена.

Справа и слева от зоны тяжелых электронов будут двигаться с большой скоростью легкие электроны. 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, а поэтому образуется область, обедненная электронами, — область положительного объемного заряда (рис. 5-11). *Слой объемного заряда с повышенной концентрацией тяжелых электронов носит название электрического домена.*

Поскольку внутри домена напряженность поля возрастает, то растет в нем и скорость движения электронов. Но так как напряжение, приложенное к образцу, не меняется, то напряженность поля вне домена, а следовательно, и скорость движения электронов падают. Через некоторый промежуток времени установится динамическое равновесие, при котором скорости движения электронов внутри домена v_d и вне его v_b будут одинаковы, т. е.

$$v_d = v_b \quad (5-137)$$

или

$$\mu_1 \mathcal{E}_d = \mu_2 \mathcal{E}_b \quad (5-138)$$

Равновесному состоянию будет соответствовать напряженность поля \mathcal{E}_d в домене и \mathcal{E}_b вне домена. При этом установившаяся скорость движения домена v_d будет меньше максимальной скорости движения электронов, которую они имеют при $\mathcal{E}_{кр}$ (рис. 5-9). Поэтому в момент подключения к образцу напряжения (через $t = t_1$) ток будет иметь максимальное значение $I_{макс}$, как это представлено на рис. 5-12. Сразу же начнется процесс образования домена, и поскольку этот процесс кратковременный, то ток практи-

чески мгновенно спадает до значения $I_{\text{мин}}$, определяемого скоростью установившегося движения домена v_d :

$$I_{\text{мин}} = sen_0 v_d \quad (5-139)$$

где s — площадь сечения образца.

Минимальное значение тока сохраняется в течение всего времени движения домена

$$T = \frac{\mathcal{L}}{v_d}. \quad (5-140)$$

По достижении конца образца (анода) домен начинает разрушаться и ток в образце возрастает. Как только величина тока в образце достигнет значения $I_{\text{макс}}$, начинается образование нового домена. При этом процесс образования домена является существенно более быстрым, чем процесс нарастания тока (рис. 5-12).

Поскольку эффект образования домена представляет собой объемную неустойчивость, он не зависит от внешних элементов схемы. Скорость движения домена не зависит от величины внешнего напряжения, и домен передвигается со скоростью дрейфа 10^7 см/сек. В результате движения домена по кристаллу во внешней цепи появляется импульс тока, скважность импульсов определяется временем T прохождения домена [формула (5-140)]. При длине образца 50 мкм, как показывает расчет, частота колебаний должна составлять около 2 Гц.

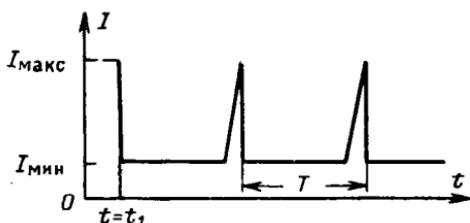


Рис. 5-12. Эпюра тока через образец арсенида галлия.

Несмотря на то что в кристалле может быть несколько неоднородностей, на которых могут формироваться домены, однако 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. Важно также, чтобы расстояние между минимумами A и B не было весьма велико, ибо для перехода электрона во второй минимум потребуется большая напряженность поля $\mathcal{E}_{\text{кр}}$, при которой возможно увеличение концентрации электронов за счет ударной ионизации или туннельного эффекта.

Ударная ионизация. В сильном электрическом поле, если оно возрастает до величины примерно 10^5 в/см, электрон (или дырка) приобретает энергию, достаточную для ионизации вещества, в ре-

зультате чего возникает электронно-дырочные пары (рис. 5-13 а, б), которые в свою очередь ускоряются и генерируют дополнительные свободные носители заряда. Этот процесс создания носителей 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 уровней, которая требует небольших полей и наблюдается при низких температурах, когда примесные

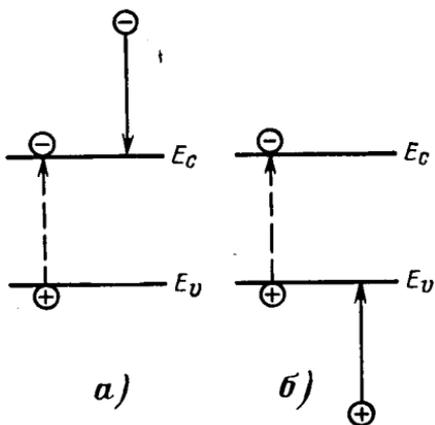


Рис. 5-13. Два возможных варианта ударной ионизации.

атомы неионизированы. Ударная ионизация примесей элементов третьей и пятой групп в германии происходит при полях приблизительно $5-10$ в/см.

Ударная ионизация характеризуется коэффициентом ударной ионизации $\alpha(E)$, который равен среднему числу ионизационных столкновений на единице длины. Эта величина определяется соотношением

$$\alpha(E) = \frac{1}{v_d} \int_0^{\infty} W_{и}(E) f(E) N(E) dE, \quad (5-141)$$

где $W_{и}(E)$ — вероятность ионизации, рассчитываемая квантовомеханическим путем; $f(E)$ — функция распределения; $N(E)$ — плотность квантовых состояний; v_d — скорость дрейфа. Для полупроводников с большой диэлектрической проницаемостью вероятность ионизации (для Ge и Si) равна:

$$W_{и}(E) \approx A (E - E_{и})^2, \quad (5-142)$$

где A — коэффициент пропорциональности, а $E_{и}$ — энергия ионизации.

При ударной ионизации участвуют три носителя, которые примерно поровну распределяют между собой энергию первоначального носителя; пороговая энергия ионизации $E_{и} = 2\Delta E_g$, где ΔE_g — ширина запрещенной зоны. Зависимость коэффициента ударной ионизации 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 $0 < E < \infty$ получить нельзя. Можно получить в явном виде два участка для $f(E)$ в области $E < E_{и}$ и $E > E_{и}$. Ускорение электронов до порога ионизации $E_{и}$ зависит от соотношения двух факторов — ускорения

во внешнем электрическом поле и рассеяния энергии при столкновениях с фононами.

Наибольший вклад в релаксацию по энергиям вносят оптические фононы. При столкновениях с оптическим фононом энергия, теряемая электроном, равна $\hbar\omega_0 = \text{const}$, а средняя длина свободного пробега $l \approx \text{const}$ и не зависит от энергии. Ускорение до порога ионизации E_n можно иллюстрировать следующей условной диаграммой (рис. 5-14). Здесь по вертикали отложена энергия, приобретаемая электроном; а по горизонтали — расстояние, проходимое им в среднем между двумя ионизирующими столкновениями. Как видно из рис. 5-14, существуют две возможности ускорения электронов до порога ионизации.

В первом случае в очень сильных электрических полях \mathcal{E} , 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,

$$\alpha(E) \sim e^{-\frac{E_n}{\mathcal{E}^2}}. \quad (5-143)$$

Здесь можно провести аналогию с диффузионным движением частицы, поэтому указанное приближение для $\alpha(E)$ называют д и ф ф у з и о н н ы м.

Во втором случае, когда поля слабые ($e\mathcal{E}l \ll \hbar\omega_0$), носители набирают энергию на расстоянии, равном многим длинам свободного пробега (линия 2). В такой ситуации существует только одна возможность набрать энергию E_n , когда носители не сталкиваются с фононами до порога ионизации. Поскольку путь до порога ионизации $L_n = E_n/e\mathcal{E}$; а средний свободный пробег между столкновениями с фононами l , то вероятность пробега до порога ионизации

без столкновений есть $e^{-\frac{L_n}{l}} \sim e^{-\frac{E_n}{e\mathcal{E}l}}$. В связи с этим функция распределения, а стало быть, и коэффициент ударной ионизации пропорциональны величине

$$\alpha(E) \sim e^{-\frac{E_n}{e\mathcal{E}l}}. \quad (5-144)$$

Функция распределения будет сильно вытянута вдоль направления электрического поля, т. е. в направлении дрейфа носителей. Поэтому это приближение для $\alpha(E)$ называют д р е й ф о в ы м.

Как показали эксперименты, в Ge и Si в основном $\alpha(E)$ определяется законом вида (5-143). Однако в узких p - n переходах (особенно четко это проявляется в Si) наблюдается зависи-

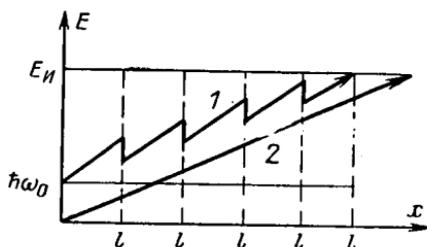


Рис. 5-14. Условная диаграмма ударной ионизации.

мость вида (5-144). Это имеет место при напряженности поля $\mathcal{E} = 5 \cdot 10^5$ в/см.

Эффект Зинера. 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. § 2-7). Наклон зон тем больше, чем больше величина напряженности электрического поля \mathcal{E} . На рис. 5-15 представлена зонная структура беспримесного полупроводника при наличии сильного электрического поля.

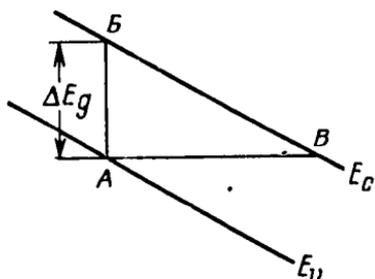


Рис. 5-15. Зонная структура беспримесного полупроводника при наличии сильного электрического поля.

В этом случае, как следует из рис. 5-15, 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 ΔE_g . Эффективная ширина барьера $AB = \Delta x$ может быть определена из разности потенциальной энергии электрона в зоне проводимости в точке B , а в валентной зоне — в точке A . Так как потенциальная энергия с точностью до аддитивной постоянной есть

$$U(B) - U(A) = -e\mathcal{E}\Delta x = -\Delta E_g, \quad (5-145)$$

то эффективная ширина барьера будет:

$$\Delta x = \frac{\Delta E_g}{e\mathcal{E}}. \quad (5-146)$$

Согласно формуле (5-146) 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 механики¹, вероятность перехода сквозь барьер треугольной формы имеет вид:

$$D = D_0 e^{-\frac{2\sqrt{m^*}(\Delta E_g)^{3/2}}{2e\hbar\mathcal{E}}}. \quad (5-147)$$

Вероятность туннельного перехода одна и та же как для переходов из валентной зоны в зону проводимости, так из зоны проводимости в валентную зону. Но поскольку в валентной зоне электро-

¹ Шпольский Э. В., Атомная физика, т. 1, Гостехиздат, 1949.

нов намного больше, чем в зоне проводимости, то будет иметь место преимущественный переход электронов из валентной зоны в зону проводимости. Туннельный эффект может приводить к значительному увеличению свободных носителей заряда. Он наблюдается в полях 10^6 в/см.

Глава шестая

ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА

6-1. РАВНОВЕСНЫЕ И НЕРАВНОВЕСНЫЕ НОСИТЕЛИ ЗАРЯДА

Рассмотрим донорный полупроводник, находящийся в состоянии термодинамического равновесия с окружающей средой. В результате тепловой генерации электроны с донорной примеси переходят в зону проводимости. При достаточно высоких температурах возможен перебор электронов из валентной зоны непосредственно в зону проводимости. На рис. 6-1 стрелками показаны переходы электронов при тепловом возбуждении. Свободные носители заряда, возникающие в результате термической генерации и находящиеся в тепловом равновесии с кристаллической решеткой, называются равновесными.

На рис. 6-1 изображены распределение плотности квантовых состояний $N(E)$ в зоне проводимости и валентной зоне, функция распределения Ферми — Дирака $f_0(E)$ и концентрации равновесных электронов n_0 и дырок p_0 , которые занимают состояния вблизи краев соответствующих зон (заштрихованные области).

Одновременно с генерацией свободных носителей идет процесс рекомбинации: электроны возвращаются в свободные состояния в валентной зоне, в результате чего исчезают свободный электрон и свободная дырка. В условиях термодинамического равновесия эти процессы полностью 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 кон-

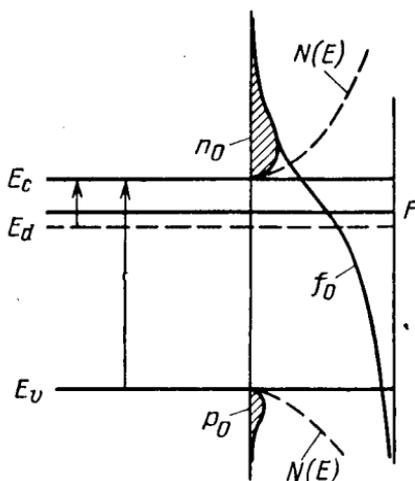


Рис. 6-1. Тепловая генерация свободных носителей и их концентрация в состоянии термодинамического равновесия.

центраций свободных носителей заряда

$$r_0 = \gamma_r n_0 p_0. \quad (6-1)$$

Здесь γ_r — коэффициент пропорциональности, называемый коэффициентом рекомбинации. Для равновесного состояния полупроводника справедливо равенство

$$g_0 = r_0, \quad (6-2)$$

которое является выражением принципа детального равновесия.

Помимо тепловой генерации, имеются другие механизмы, приводящие к возникновению свободных носителей заряда. Например, они могут образоваться при облучении полупроводника светом, за счет разрыва валентных связей в сильных электрических полях или в результате инжекции с помощью p - n перехода. Во всех этих случаях создается некоторая концентрация неравновесных свободных электронов Δn и дырок Δp , которые в момент возникновения могут иметь кинетическую энергию, значительно превышающую среднюю тепловую энергию равновесных частиц. Однако в результате рассеяния на несовершенствах кристаллической решетки носители заряда довольно быстро передают ей свою избыточную энергию. Средняя длина свободного пробега электронов имеет порядок 10^{-6} см, а тепловая скорость при комнатной температуре составляет приблизительно 10^7 см/сек. При этом среднее время между двумя столкновениями равно $\tau \approx 10^{-13}$ сек. Для рассеяния избыточной энергии порядка 1 эв неравновесные электроны должны совершить около 1000 столкновений, это значит, что уже через 10^{-10} сек они приобретут температуру кристаллической решетки и не будут отличаться от равновесных носителей заряда, поэтому распределение по энергиям неравновесных и равновесных носителей будет одинаково. Этот процесс сводится к тому, что неравновесные электроны, рассеивая свою энергию, как бы «опускаются» к нижнему краю зоны проводимости, а дырки «поднимаются» к верхнему краю валентной зоны. Если при этом концентрация неравновесных носителей мала по сравнению с равновесными, то можно считать, что энергия кристалла практически не изменяется, а следовательно, не меняется температура кристалла и концентрация равновесных носителей заряда. В этом случае общее количество электронов и дырок будет равно соответственно

$$n = n_0 + \Delta n; \quad (6-3)$$

$$p = p_0 + \Delta p.$$

Концентрацию свободных электронов для неравновесного состояния можно выразить с помощью формулы, аналогичной выражению (3-39),

$$n = n_0 + \Delta n = \int_{E_c}^{\infty} P_e(E) N(E) dE, \quad (6-4)$$

где $P_e(E)$ — функция распределения для электронов, отличная от равновесной функции распределения $f_0(E)$ и стремящаяся к ней при приближении системы к состоянию термодинамического равновесия. Это выражение можно представить в виде, аналогичном

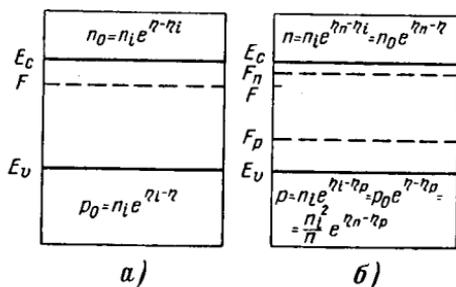


Рис. 6-2. Расщепление уровня Ферми на два квазиуровня — для электронов F_n и дырок F_p .
 а — равновесное состояние; б — неравновесное состояние.

соотношению (3-41), введя квазиуровень Ферми для электронов в неравновесном состоянии F_n так, чтобы выполнялось условие

$$n = N_c F_{1/2}(\eta_n), \quad (6-5)$$

где

$$\eta_n = \frac{F_n - E_c}{k_0 T} \quad (6-6)$$

есть приведенный квазиуровень Ферми для электронов.

Для невырожденного полупроводника с учетом (3-58), (3-60) и (3-63) можно записать:

$$n = N_c e^{\eta_n} = n_0 e^{\eta_n - \eta} = n_i e^{\eta_n - \eta_i}. \quad (6-7)$$

Аналогично концентрацию свободных дырок можно выразить следующим образом:

$$p = p_0 + \Delta p = \int_{-\infty}^{E_v} P_h(E) N(E) dE = N_v F_{1/2}(-\eta_p - \varepsilon_i), \quad (6-8)$$

где

$$\eta_p = \frac{F_p - E_c}{k_0 T} \quad (6-9)$$

есть приведенный квазиуровень Ферми для дырок.

При отсутствии вырождения, учитывая формулы (3-61) и (3-62), можно написать:

$$p = N_v e^{-\eta_p - \varepsilon_i} = p_0 e^{\eta - \eta_p} = n_i e^{\eta_i - \eta_p}. \quad (6-10)$$

Таким образом, в неравновесном состоянии уровень Ферми как бы расщепляется на два квазиуровня — для электронов F_n и для дырок F_p (рис. 6-2).

Произведение концентраций электронов и дырок для неравновесного состояния отличается теперь от его значения для равновесного состояния

$$np = n_i^2 e^{\eta_n - \eta_p} = n_0 p_0 e^{\eta_n - \eta_p}. \quad (6-11)$$

Расстояние между квазиуровнями Ферми ($\eta_n - \eta_p$) характеризует отклонение от состояния термодинамического равновесия

$$\frac{np}{n_0 p_0} = \frac{np}{n_i^2} = e^{\eta_n - \eta_p}. \quad (6-12)$$

Это соотношение выражает связь между концентрациями дырок и электронов в неравновесном состоянии.

6-2. БИПОЛЯРНАЯ СВЕТОВАЯ ГЕНЕРАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА

Рассмотрим полупроводник, в котором в результате световой инжекции создается некоторая избыточная концентрация электронов и дырок. *Генерация при образовании двух типов носителей заряда называется биполярной генерацией.* В случае биполярной световой генерации поглощение кванта света сопровождается разрывом валентной связи и количество образующихся неравновесных электронов и дырок одинаково

$$\Delta n = \Delta p. \quad (6-13)$$

Следовательно, в тонком слое у поверхности полупроводника, в котором происходит поглощение света, концентрация свободных электронов и дырок будет повышенной и определяется равенством (6-3).

Параллельно с процессом генерации идет процесс рекомбинации, и в стационарном состоянии число электронов и дырок, возбуждаемых за единицу времени в результате разрыва валентных связей, равно числу электронов и дырок, рекомбинирующих за это же время. Поскольку неравновесные носители через очень малое время становятся физически неотличимыми от равновесных, можно считать, что они будут иметь тот же коэффициент рекомбинации γ_r , что и равновесные носители.

После выключения возбуждающего света концентрации электронов и дырок будут уменьшаться в результате рекомбинации. При этом скорость убывания числа свободных носителей определяется разностью скоростей рекомбинации и тепловой генерации

$$-\left(\frac{dn}{dt}\right)_r = -\left(\frac{dp}{dt}\right)_r = \gamma_r np - g_0. \quad (6-14)$$

Принимая во внимание соотношения (6-1) и (6-2), равенство (6-14) можно переписать следующим образом:

$$-\left(\frac{dn}{dt}\right)_r = \gamma_r (np - n_0 p_0) = \gamma_r (n_0 \Delta p + p_0 \Delta n + \Delta n \Delta p). \quad (6-15)$$

Учитывая, что в случае биполярной световой генерации при малом уровне инжекции $\Delta n = \Delta p \ll (n_0 + p_0)$, получаем:

$$\left(\frac{dn}{dt}\right)_r = -\gamma_r (n_0 + p_0) \Delta n = -\gamma_r (n_0 + p_0) (n - n_0). \quad (6-16)$$

Введем обозначение

$$\tau = \frac{1}{\gamma_r (n_0 + p_0)}. \quad (6-17)$$

Величина τ , имеющая размерность времени, называется в ре-
менем жизни неравновесных носителей
заряда.

Уравнение (6-16) примет теперь такой вид:

$$\left(\frac{dn}{dt}\right)_r = -\frac{n - n_0}{\tau} = -\frac{\Delta n}{\tau}, \quad (6-18)$$

откуда

$$\Delta n = \Delta n(0) e^{-\frac{t}{\tau}}. \quad (6-19)$$

В случае малого уровня биполярной инжекции концентрация неравновесных носителей заряда после прекращения возбуждения уменьшается по экспоненциальному закону, и за время τ число их в результате рекомбинации убывает в e раз. Следовательно, τ представляет собой среднее время существования избыточной концентрации носителей заряда. Для собственного полупроводника скорости убывания числа электронов и дырок равны, и величина τ определяет время жизни электронно-дырочных пар. Значение объемного времени жизни в зависимости от типа полупроводникового материала и от степени его очистки может изменяться в очень широких пределах от 10^{-2} до 10^{-8} сек. Поскольку величина τ достаточно велика по сравнению со временем релаксации, свободные электроны и дырки, диффундируя из области, где их концентрация повышена, в глубь полупроводника, за время своей жизни успеют пройти большое расстояние. Следовательно, в случае биполярной генерации возникновение носителей происходит в одной области образца, а рекомбинация их осуществляется на значительном расстоянии в другой части объема полупроводника.

В случае большого уровня биполярной генерации, когда $\Delta n \gg (n_0 + p_0)$, из уравнения (6-15) найдем, что

$$-\left(\frac{dn}{dt}\right)_r = \gamma_r (\Delta n)^2. \quad (6-20)$$

Разделяя переменные, запишем:

$$\frac{dn}{(\Delta n)^2} = -\gamma_r dt, \quad (6-21)$$

откуда

$$\Delta n = \frac{\Delta n(0)}{1 + \gamma_r \Delta n(0) t}, \quad (6-22)$$

где $\Delta n(0)$ — неравновесная концентрация электронов в момент выключения возбуждающего света.

Из равенства (6-22) следует, что при квадратичной рекомбинации концентрация неравновесных носителей уменьшается по гиперболическому закону.

Если ввести время жизни, то для квадратичной рекомбинации согласно уравнению (6-20)

$$\left(\frac{dn}{dt}\right)_r = -\frac{\Delta n}{\tau_{\text{мгн}}}, \quad (6-23)$$

где

$$\tau_{\text{мгн}} = \frac{1}{\gamma_r \Delta n} \quad (6-24)$$

зависит от концентрации неравновесных носителей и в связи с этим будет переменной величиной. Однако в каждый момент времени $\tau_{\text{мгн}}$ имеет определенный смысл, являясь м г н о в е н н ы м в р е м е н е м ж и з н и. В силу этого при большом уровне инжекции имеем дело с мгновенным временем жизни, которое согласно уравнению (6-23) равно:

$$\tau_{\text{мгн}} = -\frac{\Delta n}{\frac{d}{dt}(\Delta n)}. \quad (6-25)$$

6.3. МОНОПОЛЯРНАЯ СВЕТОВАЯ ГЕНЕРАЦИЯ. 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 атомы примеси при данной температуре не ионизированы полностью. Это означает, что примесный уровень лежит достаточно глубоко в запрещенной зоне ($E_d \gg k_0 T$). Пусть облучение полупроводника светом приводит к перебросу электронов с донорных уровней в зону проводимости за счет энергии квантов света. В этом случае имеет место *монополярная световая генерация, которая характеризуется образованием неравновесных носителей заряда одного знака* (в данном случае электронов). При этом электронейтральность полупроводника не нарушается, так как избыточный заряд свободных электронов скомпенсирован зарядом образовавшихся положительных ионов донорной примеси.

Если коэффициент поглощения света велик, неравновесные электроны генерируются лишь в тонком приповерхностном слое и у поверхности полупроводника концентрация электронов будет повышенной и равной $n = n_0 + \Delta n$, а в глубине образца n_0 . Неравновесные носители будут перемещаться в глубь полупроводника, что приведет к возникновению в его объеме отрицательного заряда по отношению к поверхности, где сосредоточены положительные

ионы. Электрическое поле, порожденное разделением зарядов, вызывает обратный поток электронов к поверхности, который уравновешивает диффузионный ток и приводит практически к полному уничтожению объемного заряда.

Изменение пространственного заряда ρ в результате протекания тока, плотность которого равна j , как известно, подчиняется уравнению непрерывности электрического заряда

$$\operatorname{div} j = \operatorname{div} (\sigma \mathcal{E}) = - \frac{\partial \rho}{\partial t}, \quad (6-26)$$

где \mathcal{E} — напряженность электрического поля.

Воспользуемся, кроме того, уравнением Пуассона

$$\operatorname{div} \mathcal{E} = \frac{\rho}{\epsilon \epsilon_0}. \quad (6-27)$$

Здесь ϵ — величина диэлектрической проницаемости полупроводника; ϵ_0 — диэлектрическая постоянная.

Решая совместно уравнение (6-26) и (2-27), получаем:

$$\frac{\partial \rho}{\partial t} = - \frac{\sigma}{\epsilon \epsilon_0} \rho. \quad (6-28)$$

Отсюда находим закон изменения объемного заряда во времени

$$\rho = \rho_0 e^{-\frac{t}{\tau_\mu}}, \quad (6-29)$$

где ρ_0 — величина заряда в момент $t = 0$;

$$\tau_\mu = \frac{\epsilon \epsilon_0}{\sigma} \quad (6-30)$$

есть диэлектрическое или максвелловское время релаксации.

Как следует из равенства (6-29), в случае монополярной генерации возникает объемный заряд, который со временем уменьшается по экспоненциальному закону с постоянной времени τ_μ . Поскольку максвелловское время релаксации τ_μ для полупроводников достаточно мало, электронное облако неравновесных носителей не может сместиться на значительное расстояние по отношению к ионам примеси и повышенная концентрация неравновесных носителей будет в той области, где происходит их генерация. Следовательно, монополярная генерация и рекомбинация носителей имеют место в одной и той же области примесного полупроводника.

Оценим порядок величины максвелловского времени релаксации. Для германия, имеющего $\sigma = 10^2 \text{ сим} \cdot \text{м}^{-1}$ и $\epsilon = 16$, $\tau_\mu = 10^{-12} \text{ сек}$.

6-4. ВИДЫ РЕКОМБИНАЦИИ

В зависимости от механизма различают три вида рекомбинации: межзонная рекомбинация, рекомбинация через локальные центры и поверхностная рекомбинация.

Межзонная рекомбинация осуществляется при переходе свободного электрона из зоны проводимости в валентную зону, что сопровождается уничтожением свободного электрона и свободной дырки. Этот процесс совершается при соблюдении законов сохранения энергии и импульса. Если энергия и квазиимпульс электрона в зоне проводимости (до рекомбинации) составляли E' и p' , а после рекомбинации они равны соответственно E и p , должны выполняться законы сохранения энергии и квазиимпульса:

$$E' = E + \Delta E; \quad p' = p + Q. \quad (6-31)$$

Здесь Q — импульс, передаваемый кристаллической решетке, а ΔE — энергия, которая выделяется при процессе рекомбинации.

В зависимости от того, каким образом расходуется энергия ΔE , межзонная рекомбинация разделяется на три вида:

если энергия, освобождающаяся в процессе рекомбинации, излучается в виде кванта света, мы имеем дело с излучательной (фотонной) рекомбинацией;

если энергия свободного электрона расходуется в конечном счете на образование фононов, рекомбинация называется безызлучательной (фононной);

в акте рекомбинации могут принимать участие три носителя; выделяющаяся энергия передается свободному электрону или дырке.

Этот тип рекомбинации называется ударной рекомбинацией или рекомбинацией Оже.

Если в полупроводниковом кристалле имеются дефекты, энергетические уровни которых находятся в запрещенной зоне, наряду с межзонной рекомбинацией идет процесс рекомбинации через локальные центры. Механизм этого процесса заключается в том, что нейтральный центр (ловушка) может захватить, например, электрон из зоны проводимости, который через некоторое время перейдет в валентную зону (что эквивалентно захвату отрицательно заряженной ловушкой дырки из валентной зоны). В соответствии со способом расходования энергии при переходе электрона на более низкий энергетический уровень рекомбинация через центры захвата также может быть излучательной, фононной или ударной.

Рекомбинация носителей заряда, протекающая на поверхности полупроводника, называется поверхностной рекомбинацией. Поверхностная рекомбинация происходит через локальные поверхностные уровни.

6-5. МЕЖЗОННАЯ ИЗЛУЧАТЕЛЬНАЯ РЕКОМБИНАЦИЯ

При межзонной излучательной рекомбинации свободный электрон из зоны проводимости переходит в валентную зону, при этом излучается квант света, энергия которого равна величине ширины запрещенной зоны:

$$h\nu = \Delta E_g. \quad (6-32)$$

При термодинамическом равновесии количество носителей, рекомбинирующих в единице объема за единицу времени с испусканием фотонов, в соответствии с принципом детального равновесия равно числу носителей, 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.

Интенсивность процесса излучательной рекомбинации пропорциональна произведению концентраций электронов и дырок

$$R_n^0 = R_p^0 = \gamma_r n_0 p_0 = \gamma_r n_i^2. \quad (6-33)$$

Здесь γ_r — коэффициент межзонной излучательной рекомбинации.

При отклонении от равновесного состояния концентрации свободных носителей заряда изменяются и равны:

$$n = n_0 + \Delta n; \quad p = p_0 + \Delta p.$$

Как отмечалось в § 6-2, коэффициент рекомбинации неравновесных носителей имеет ту же величину, что и у равновесных. При этом, если не наступает вырождения, сохраняется пропорциональность скорости рекомбинации R произведению концентраций электронов и дырок. Поэтому, определяя γ_r из соотношения (6-33), можно написать:

$$R = \gamma_r n p = R_n^0 \frac{np}{n_i^2}. \quad (6-34)$$

Определим время жизни неравновесных носителей заряда при межзонной излучательной рекомбинации. В соответствии с уравнением (6-18) для электронов имеем:

$$\tau_{nr} = - \frac{\Delta n}{\frac{dn}{dt}}. \quad (6-35)$$

Когда внешнее возбуждение снято, скорость изменения концентрации свободных электронов определяется разностью интенсивностей рекомбинации и равновесной генерации, которая равна:

$$- \left(\frac{dn}{dt} \right) = R - R_n^0. \quad (6-36)$$

С учетом этого из равенства (6-35) получим:

$$\tau_{nr} = \frac{\Delta n}{R - R_n^0} = \frac{\Delta n}{\Delta R}, \quad (6-37)$$

где ΔR — изменение скорости рекомбинации при отклонении от равновесного состояния. Аналогичное равенство можно записать для дырок

$$\tau_{pr} = \frac{\Delta p}{R - R_p^0} = \frac{\Delta p}{\Delta R}. \quad (6-38)$$

Величину ΔR определим с помощью формул (6-34) и (6-33)

$$\Delta R = R - R_{и}^0 = R_{и}^0 \frac{n\rho - n_i^2}{n_i^2}. \quad (6-39)$$

Подставив в формулу (6-39) значения n и ρ и заменив n_i^2 на $n_0\rho_0$, получим:

$$\Delta R = R_{и}^0 \frac{n_0\Delta\rho + \rho_0\Delta n + \Delta n\Delta\rho}{n_0\rho_0}. \quad (6-40)$$

Теперь выражение (6-37) принимает вид:

$$\tau_{nr} = \Delta n \frac{n_0\rho_0}{R_{и}^0 (n_0\Delta\rho + \rho_0\Delta n + \Delta n\Delta\rho)}. \quad (6-41)$$

Если полупроводник не содержит ловушек, можно считать, что $\Delta n = \Delta\rho$, при этом формула (6-41) упрощается

$$\tau_{nr} = \tau_{pr} = \frac{n_0\rho_0}{R_{и}^0 (n_0 + \rho_0 + \Delta n)}. \quad (6-42)$$

Проведем анализ полученного выражения для времени жизни при межзонной излучательной рекомбинации. В случае малого уровня инжекции $\Delta n \ll (n_0 + \rho_0)$ с учетом формулы (6-33) получим:

$$\tau_{nr} = \tau_{pr} = \frac{n_i^2}{R_{и}^0 (n_0 + \rho_0)} = \frac{1}{\gamma_r (n_0 + \rho_0)}, \quad (6-43)$$

а для собственного полупроводника, у которого $n_0 = \rho_0 = n_i$,

$$\tau_{ir} = \frac{n_i}{2R_{и}^0} = \frac{1}{2\gamma_r n_i} = \frac{1}{2\gamma_r (N_c N_v)^{1/2}} e^{\frac{\Delta E_g}{2k_0 T}}. \quad (6-44)$$

Для материала n -типа ($n_0 \gg \rho_0$) имеем:

$$\tau_r^n = \frac{\rho_0}{R_{и}^0} = 2 \frac{n_i}{n_0} \tau_{ir} = \frac{1}{\gamma_r n_0}. \quad (6-45)$$

Аналогично этому в случае дырочного полупроводника ($\rho_0 \gg n_0$) время жизни равно:

$$\tau_r^p = \frac{n_0}{R_{и}^0} = 2 \frac{n_i}{\rho_0} \tau_{ir} = \frac{1}{\gamma_r \rho_0}. \quad (6-46)$$

Анализ формулы (6-44) показывает, что время жизни неравновесных носителей в собственном полупроводнике тем меньше, чем выше температура и чем меньше ширина запрещенной зоны. Эти закономерности подтверждаются экспериментальными данными. Из формул (6-45) и (6-46) видно, что время жизни неравновесных носителей заряда в примесном полупроводнике меньше, чем в собственном, и с ростом температуры оно уменьшается.

Зависимость времени жизни для излучательной межзонной рекомбинации от степени легирования, т. е. концентрации носителей заряда при постоянной температуре, представлена на рис. 6-3.

Логарифмическую шкалу концентраций на рис. 6-3 можно рассматривать так же, как линейную шкалу положения уровня Ферми,

так как $n_0 = N_c e^{-\frac{E_c - F}{k_0 T}}$, причем ее средняя точка соответствует значению уровня Ферми для собственного полупроводника. Из

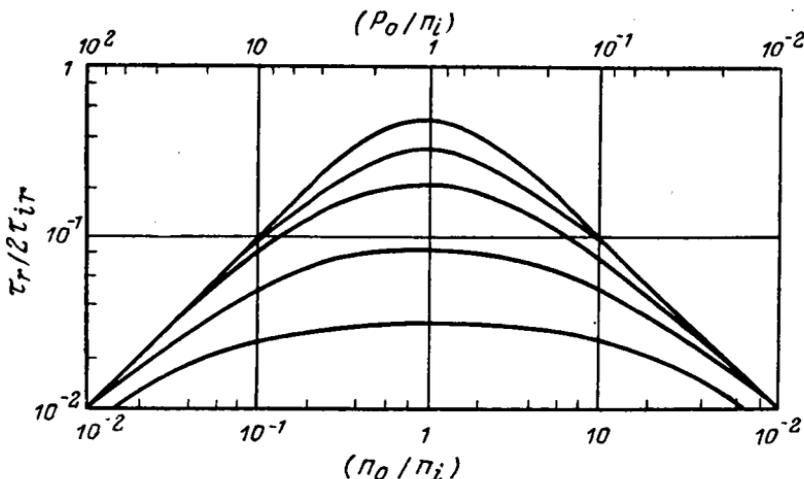


Рис. 6-3. Зависимость времени жизни для межзонной излучательной рекомбинации от концентрации носителей заряда при заданной температуре в случае низкого уровня возбуждения. В максимумах отношение $\Delta n / n_i = 0$ для верхней кривой, и соответственно равно 1, 3, 10, 30 для последующих кривых.

приведенных зависимостей видно, что с увеличением уровня инжекции время жизни в собственном полупроводнике резко снижается, а в примесном изменяется сравнительно слабо.

Рассчитаем величину R_{ii}^0 , используя для этого теорию Эйнштейна для равновесного излучения. Число квантовых состояний для фотонов в интервале импульса между P и $P + dP$ составляет:

$$\frac{P^2 dP}{h^3}. \quad (6-47)$$

Чтобы определить количество ячеек, занятых фотонами, эту величину следует умножить на функцию распределения Бозе — Эйнштейна

$$f_\Phi = \frac{1}{e^{\frac{h\nu}{k_0 T}} - 1}. \quad (6-48)$$

Общее число фотонов с импульсом P , излучаемых в телесный угол $d\Omega$, составит:

$$dn_\Phi = \frac{2}{h^3} f_\Phi P^2 dP. \quad (6-49)$$

Коэффициент 2 учитывает различную поляризацию света.

Импульс фотона связан с энергией ($E = h\nu$) следующим образом:

$$P = \frac{E}{v_\Phi} = \frac{h\nu}{v_\Phi} = \frac{nh\nu}{c}. \quad (6-50)$$

Здесь $v_\Phi = c/\bar{n}$ — скорость фотона в данной среде; \bar{n} — показатель преломления среды.

Из формулы (6-50) получаем:

$$dP = \frac{\bar{n}h}{c} d\nu, \quad (6-51)$$

что при подстановке в уравнение (6-49) после интегрирования по всем углам от 0 до 4π дает:

$$dN = \int_0^{4\pi} \frac{2}{h^3} f_\Phi P^2 dP d\Omega = \frac{8\pi}{h^3} f_\Phi P^2 dP = \frac{8\pi}{c^3} \frac{n^3 v^2 d\nu}{e^{\frac{h\nu}{k_0 T}} - 1}. \quad (6-52)$$

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:

$$R_n^0 = R_n^i = \int_0^\infty g(\nu) dN = \frac{8\pi}{c^3} \int_0^\infty \frac{\bar{n}^3(\nu) g(\nu) \nu^2 d\nu}{e^{\frac{h\nu}{k_0 T}} - 1}. \quad (6-53)$$

Время жизни фотона τ_Φ есть величина, обратная вероятности его поглощения. Величину τ_Φ можно определить, если известны длина свободного пробега l_Φ и скорость фотона v_Φ . В то же время длина свободного пробега фотона обратно пропорциональна коэффициенту поглощения $\alpha(\nu)$, т. е. $l_\Phi = 1/\alpha$. В результате будем иметь:

$$\tau_\Phi = \frac{1}{g(\nu)} = \frac{l_\Phi}{v_\Phi} = \frac{\bar{n}}{c\alpha(\nu)}. \quad (6-54)$$

Подставляя значение $g(\nu)$ из формулы (6-54) в (6-53), находим:

$$R_n^0 = \gamma_r n_i^2 = \frac{8\pi}{c^2} \int_0^\infty \frac{\bar{n}^2(\nu) \alpha(\nu) \nu^2 d\nu}{e^{\frac{h\nu}{k_0 T}} - 1}. \quad (6-55)$$

Величины R_n^0 и значения времени жизни для собственного полупроводника $\tau_{ir} = n_i/2R_n^0$, рассчитанные для некоторых материалов, представлены в табл. 6-1. Там же приведены экспериментальные данные для этих полупроводников.

Из приведенных данных табл. 6-1 следует, что с уменьшением ширины запрещенной зоны полупроводника величина времени жизни для межзонной излучательной рекомбинации снижается. Рассчитанные значения τ_{ir} 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 означает, что для этих материалов велика вероятность безызлучательных переходов.

Таблица 6-1

Экспериментальные данные по параметрам рекомбинации некоторых полупроводников

Вещество	Ширина запрещенной зоны, эВ	n_i , $см^{-3}$	R_{ii}^0 , $см^{-3} \cdot сек^{-1}$	τ_{ir} , сек	$\tau_{экс}$, сек
Si	0,1	$1,4 \cdot 10^{10}$	$2 \cdot 10^{19}$	3,5	10^{-3}
Ge	0,7	$2,4 \cdot 10^{13}$	$3,7 \cdot 10^{13}$	0,3	10^{-3}
PbS	0,37	$3 \cdot 10^{15}$	$14 \cdot 10^{20}$	10^{-5}	$9 \cdot 10^{-6}$
PbSe	0,22	$2 \cdot 10^{17}$	$3,3 \cdot 10^{22}$	$3 \cdot 10^{-6}$	—
PbTe	0,27	$6 \cdot 10^{16}$	$1,8 \cdot 10^{22}$	$1,7 \cdot 10^{-6}$	—
InSb	0,18	$2,2 \cdot 10^{16}$	$2,6 \cdot 10^{22}$	$0,4 \cdot 10^{-6}$	$0,1 \cdot 10^{-6}$

Согласно формулам (6-45) и (6-46) при увеличении степени легирования время жизни неравновесных носителей заряда снижается. Например, для сильно легированного германия ($n_0 \approx 10^{19} см^{-3}$) оно составляет:

$$\tau_r^n = 2 \frac{n_i}{n_0} \tau_{ir} \approx 2 \frac{2,4 \cdot 10^{13}}{10^{19}} 0,3 \approx 10^{-6} \text{ сек.}$$

Казалось бы, с ростом степени легирования материала должна увеличиваться вероятность межзонной излучательной рекомбинации. Однако с повышением концентрации примеси появляется большее количество дефектов и начинает преобладать рекомбинация через локальные центры. Поэтому для исследования процессов межзонной излучательной рекомбинации следует использовать наиболее чистые материалы.

6-6. МЕЖЗОННАЯ УДАРНАЯ РЕКОМБИНАЦИЯ

Если происходит столкновение одновременно трех свободных носителей заряда, может иметь место рекомбинация двух из них с передачей энергии третьему носителю, который переходит на более высокий уровень энергии в зоне. При этом мы имеем дело с явлением, обратным процессу ударной ионизации, при котором за счет перехода свободной частицы в зону на более низкий энергетический уровень происходит возникновение пары электрон-дырка. На рис. 6-4 схематично изображены процессы ударной рекомбинации с участием в качестве третьего носителя электрона (а) или дырки (б).

Вероятность столкновения пары электрон-дырка со свободным электроном пропорциональна $n^2 p$, а с дыркой — $p^2 n$. Таким образом, уменьшение концентрации свободных электронов или дырок

в результате межзонной ударной рекомбинации определяется выражением

$$-\frac{dn}{dt} = -\frac{dp}{dt} = \eta^{(n)}n^2p + \eta^{(p)}np^2 - \eta^{(n)}n_0^2p_0 - \eta^{(p)}n_0p_0^2, \quad (6-56)$$

где $\eta^{(n)}$ и $\eta^{(p)}$ — коэффициенты ударной рекомбинации с участием в качестве третьего носителя электрона или дырки. Два последних члена в выражении (6-56) учитывают интенсивность генерации при термодинамическом равновесии. Следует отметить, что коэффициенты ударной рекомбинации отличаются по размерности от соответствующих коэффициентов в случае рекомбинации другого типа.

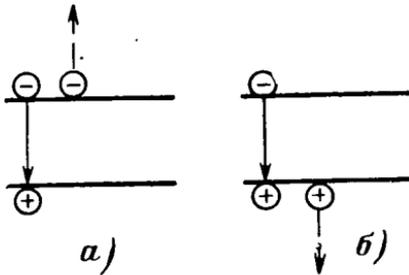


Рис. 6-4. Схема процессов ударной рекомбинации.

Для полупроводника, не содержащего значительного количества ловушек, можно принять, что $\Delta n = \Delta p$. Подставив в выражение (6-56) значения n и p и расположив члены уравнения по степеням Δn , получим:

$$-\frac{d}{dt} \Delta n = -\frac{d}{dt} \Delta p = \Delta n (\eta^{(n)}n_0^2 + 2\eta^{(n)}n_0p_0 + \eta^{(p)}p_0^2 + 2\eta^{(p)}n_0p_0) + \Delta n^2 (\eta^{(n)}p_0 + 2\eta^{(n)}n_0 + 2\eta^{(p)}p_0 + \eta^{(p)}n_0) + \Delta n^3 (\eta^{(n)} + \eta^{(p)}). \quad (6-57)$$

Найдем выражение для времени жизни при межзонной ударной рекомбинации согласно (6-18), считая, что уровень инжекции мал ($\Delta n \ll n_0 + p_0$). Для этого отбросим в (6-57) члены, содержащие Δn во второй и третьей степени, и будем иметь:

$$\tau_{nu} = \tau_{pu} = -\frac{\Delta n}{\frac{d}{dt} \Delta n} = \frac{1}{\eta^{(n)}n_0^2 + 2\eta^{(n)}n_0p_0 + \eta^{(p)}p_0^2 + \eta^{(p)}n_0p_0}. \quad (6-58)$$

Рассмотрим полученную формулу для различных частных случаев.

1. Собственный полупроводник ($n_0 = p_0 = n_i$)

$$\tau_{iu} = \frac{1}{3(\eta^{(n)} + \eta^{(p)})n_i^2} = \frac{1}{3(\eta^{(n)} + \eta^{(p)})N_cN_v} e^{\frac{\Delta E_g}{k_0T}}. \quad (6-59)$$

Величина τ_{iu} снижается с ростом температуры и с уменьшением ширины запрещенной зоны.

2. Донорный полупроводник ($n_0 \gg p_0; n_0 \gg n_i$)

$$\tau_u^n = \frac{1}{\eta^{(n)}n_0^2}. \quad (6-60)$$

3. Акцепторный полупроводник ($p_0 \gg n_0, p_0 \gg n_i$)

$$\tau_u^p = \frac{1}{\eta^{(p)}p_0^2}. \quad (6-61)$$

Из полученных выражений следует, что время жизни для собственного полупроводника при 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 и время жизни снижается. Зависимость τ_u от положения уровня Ферми приведена на рис. 6-5.

Как видно из рис. 6-5, время жизни для собственного полупроводника максимально. Оно уменьшается с ростом степени легирования.

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, 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 уравнении (6-57) можно отбросить два первых слагаемых, т. е. члены, содержащие Δn и Δn^2 . В результате будем иметь:

$$-\frac{d}{dt} \Delta n = -\frac{d}{dt} \Delta p = \Delta n^3 (\eta^{(n)} + \eta^{(p)}). \quad (6-62)$$

Мгновенное время жизни при этом равно:

$$\tau_{\text{мгн}} = \frac{\Delta n}{-\frac{d}{dt} \Delta n} = \frac{1}{(\eta^{(n)} + \eta^{(p)}) \Delta n^2}. \quad (6-63)$$

Из (6-63) следует, что мгновенное время жизни для ударной межзонной рекомбинации при высоком уровне возбуждения очень сильно зависит от концентрации неравновесных носителей заряда.

6-7. РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА ЧЕРЕЗ ЦЕНТРЫ ЗАХВАТА

Дефекты, которые могут захватывать свободные носители заряда с последующим их освобождением, называются центрами захвата или ловушками. Теория рекомбинации через центры захвата (ловушки), наиболее часто реализующаяся на практике, была разработана Шокли и Ридом. Реальные полупроводниковые материалы содержат обычно примеси нескольких типов, каждая из которых может создавать один или несколько уровней в запрещенной зоне полупроводника. Мы ограничимся рассмотрением простейшего случая, когда в полупроводнике имеется один тип ловушек, создающих единственный энергетический уровень, причем центры захвата

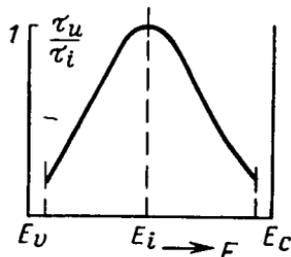


Рис. 6-5. Время жизни при ударной межзонной рекомбинации в зависимости от положения уровня Ферми.

могут быть либо нейтральны, либо однократно заряжены. Такие локальные центры называют простыми.

Пусть концентрация ловушек равна N_t , а их энергетический уровень E_t лежит вблизи центра запрещенной зоны. Процесс рекомбинации через локальные центры проиллюстрирован на рис. 6-6.

В полупроводнике донорного типа, имеющем большое число свободных электронов, рекомбинация проходит следующим образом. Нейтральная ловушка захватывает электрон из зоны проводимости, приобретая отрицательный заряд (переход 1). Затем электрон с уровня ловушки переходит в валентную зону, что эквивалентно захвату отрицательно заряженной ловушки дырки (переход 2).

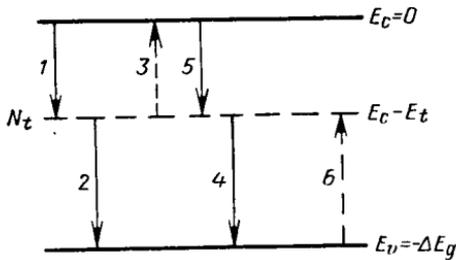


Рис. 6-6. Переходы электронов, связанные с рекомбинацией через локальные центры.

центра рекомбинации в валентную зону (переход 4), а затем происходит захват электрона из зоны проводимости (переход 5). При малой концентрации электронов возможен возврат дырки в валентную зону (переход 6). *Дефект, захватывающий последовательно электрон и дырку и приводящий, таким образом, к рекомбинации пары свободных носителей заряда, называется центром рекомбинации или рекомбинационной ловушкой.*

Перейдем к количественному рассмотрению описанных процессов. Пусть энергетический уровень ловушки лежит между центром запрещенной зоны и дном зоны проводимости. Интенсивность захвата ловушкой электронов $R_{зах}^n$ пропорциональна их концентрации в зоне проводимости n и числу свободных мест на уровне ловушек $N_t(1 - f_t)$:

$$R_{зах}^n = \gamma_n n N_t (1 - f_t), \quad (6-64)$$

где γ_n — коэффициент захвата электронов; $f_t = f(E_c - E_t)$ — вероятность того, что электрон находится на уровне ловушки.

Количество электронов, возвращающихся в зону проводимости $R_{воз}^n$, пропорционально их концентрации на ловушках $f_t N_t$:

$$R_{воз}^n = \beta_n f_t N_t, \quad (6-65)$$

где β_n — коэффициент ионизации электронов.

Изменение концентрации электронов в зоне проводимости определяется разницей интенсивностей процессов захвата и ионизации

$$-\frac{dn}{dt} = \gamma_n(1 - f_t) n N_t - \beta_n f_t N_t. \quad (6-66)$$

Чтобы определить связь между коэффициентами γ_n и β_n , рассмотрим состояние термодинамического равновесия ($dn/dt = 0$). Поскольку неравновесные носители большую часть своего времени жизни физически неотличимы от равновесных, они характеризуются теми же коэффициентами захвата и ионизации (мы рассматриваем случай, когда в неравновесном состоянии не наступает вырождения). Учитывая это, из равенства (6-66) получаем:

$$\beta_n = \gamma_n n_0 \left(\frac{1}{f_{0t}} - 1 \right) = \gamma_n N_c e^{\frac{F - E_c}{k_0 T}} e^{\frac{E_c - E_t - F}{k_0 T}} = \gamma_n N_c e^{-\frac{E_t}{k_0 T}}. \quad (6-67)$$

При выводе этого соотношения мы использовали тот факт, что в состоянии термодинамического равновесия вероятность заполнения уровня ловушки электронами определяется выражением

$$f_{0t} = f_0(E_c - E_t) = \frac{1}{e^{\frac{E_c - E_t - F}{k_0 T}} + 1}.$$

Введем обозначение

$$N_c e^{-\frac{E_t}{k_0 T}} = n_1. \quad (6-68)$$

Величина n_1 равна равновесной концентрации электронов в зоне проводимости, когда уровень Ферми совпадает с уровнем ловушки.

Используя равенства (6-67) и (6-68), преобразуем уравнение (6-66) следующим образом:

$$-\frac{dn}{dt} = \gamma_n N_t [(1 - f_t) n - n_1 f_t]. \quad (6-69)$$

Это выражение определяет скорость изменения концентрации электронов в зоне проводимости в результате процессов захвата электронов ловушками и генерации их с ловушек.

Скорость изменения концентрации дырок в валентной зоне можно представить так:

$$-\frac{dp}{dt} = \gamma_p N_t f_t p - \beta_p N_t (1 - f_t). \quad (6-70)$$

Здесь первое слагаемое определяет количество дырок, захватываемых ловушками, а второе — число дырок, возвращающихся в валентную зону. Величины γ_p и β_p представляют собой соответственно коэффициенты захвата и ионизации для дырок.

Из условия термодинамического равновесия ($dp/dt = 0$) найдем:

$$\beta_p = \gamma_p p_0 \left(\frac{1}{1/f_{0t}} - 1 \right) = \gamma_p N_v e^{\frac{E_t - \Delta E_g}{k_0 T}} = \gamma_p p_1, \quad (6-71)$$

откуда найдем, что

$$p_1 = N_v e^{\frac{E_t - \Delta E_g}{k_0 T}}, \quad (6-72)$$

есть равновесная концентрация дырок в валентной зоне, когда уровень Ферми совпадает с уровнем ловушек.

Подставляя выражение (6-71) в формулу (6-70), получаем:

$$-\frac{dp}{dt} = \gamma_p N_t [pf_t - p_1(1 - f_t)]. \quad (6-73)$$

Это выражение определяет число электронов, уходящих с уровней ловушек в валентную зону.

Используя уравнения (6-69) и (6-73), находим количество электронов, находящихся на ловушках:

$$N_t \frac{df_t}{dt} = \left(-\frac{dn}{dt}\right) - \left(-\frac{dp}{dt}\right) = \gamma_p N_t [pf_t - p_1(1 - f_t)] - \gamma_n N_t [n(1 - f_t) - n_1 f_t]. \quad (6-74)$$

Напишем условие электронейтральности для полупроводника, содержащего некоторое количество ионов донорной примеси N_d^+ . Для равновесного случая условие электронейтральности имеет вид:

$$n_0 + N_d f_{0t} = p_0 + N_d^+. \quad (6-75)$$

Будем считать, что число ионов донорной примеси не меняется и в неравновесном состоянии. Тогда

$$n + N_d f_t = p + N_d^+. \quad (6-76)$$

Вычитая выражение (6-76) из уравнения (6-75), находим соотношение для концентрации неравновесных носителей

$$n - n_0 + N_t (f_t - f_{0t}) = p - p_0$$

или

$$\Delta n + N_t \Delta f_t = \Delta p. \quad (6-77)$$

Если концентрация ловушек в данном материале мала по сравнению с плотностью неравновесных носителей, т. е. $N_t \ll \Delta n$, то количеством электронов $N_t \Delta f_t$, оседающих на ловушках, можно пренебречь. При этом условии неравновесная концентрация электронов равна неравновесной концентрации дырок, т. е.

$$\Delta n = \Delta p. \quad (6-78)$$

В этом случае функцию распределения для ловушек можно найти из выражения (6-74), приравняв нулю его левую часть:

$$f_t = \frac{\gamma_n n_1 + \gamma_p p_1}{\gamma_n (n + n_1) + \gamma_p (p + p_1)}. \quad (6-79)$$

Подставив найденное значение для f_i в выражение (6-69), получим:

$$-\frac{dn}{dt} = \frac{\gamma_n \gamma_p N_t (np - n_1 p_1)}{\gamma_n (n + n_1) + \gamma_p (p + p_1)}. \quad (6-80)$$

Отсюда определим время жизни неравновесных носителей заряда при рекомбинации через локальные центры, учитывая, что

$$n_1 p_1 = n_i^2 = n_0 p_0$$

$$\begin{aligned} \tau &= -\frac{dn}{\frac{dn}{dt}} = \frac{\gamma_n (n_0 + n_1 + \Delta n) + \gamma_p (p_0 + p_1 + \Delta p)}{\gamma_n \gamma_p N_t (n_0 + p_0 + \Delta n)} = \\ &= \frac{1}{\gamma_p N_t} \frac{n_0 + n_1 + \Delta n}{n_0 + p_0 + \Delta n} + \frac{1}{\gamma_n N_t} \frac{p_0 + p_1 + \Delta n}{n_0 + p_0 + \Delta n}. \end{aligned} \quad (6-81)$$

Введем обозначения:

$$\tau_{p0} = \frac{1}{\gamma_p N_t}; \quad \tau_{n0} = \frac{1}{\gamma_n N_t}. \quad (6-82)$$

С учетом равенства (6-82) формула (6-81) может быть записана в виде

$$\tau = \tau_{p0} = \frac{n_0 + n_1 + \Delta n}{n_0 + p_0 + \Delta n} + \tau_{n0} \frac{p_0 + p_1 + \Delta n}{n_0 + p_0 + \Delta n}. \quad (6-83)$$

Рассмотрим более подробно зависимость времени жизни от положения уровня Ферми при постоянной температуре, т. е. от степени легирования материала. В случае малого уровня инжекции, когда ($\Delta n \ll n_0 + p_0$), формула (6-83) существенно упрощается

$$\tau = \tau_{p0} \frac{n_0 + n_1}{n_0 + p_0} + \tau_{n0} \frac{p_0 + p_1}{n_0 + p_0}. \quad (6-84)$$

Итак, при малом уровне инжекции время жизни не зависит от концентрации неравновесных носителей, а определяется значениями равновесных концентраций электронов и дырок, а также положением энергетического уровня ловушки, задающим величины n_1 и p_1 .

На кривой, изображающей зависимость $\ln \tau$ от положения уровня Ферми (рис. 6-7), можно выделить четыре основные области.

Первая область. Сильно легированный полупроводник n -типа. Уровень Ферми расположен ниже E_c , но выше энергетического уровня ловушек, т. е. $(E_c - E_t) < F < E_c$. При этом справедливы следующие соотношения:

$$n_0 \gg p_0; \quad n_0 \gg n_1; \quad n_0 \gg p_1.$$

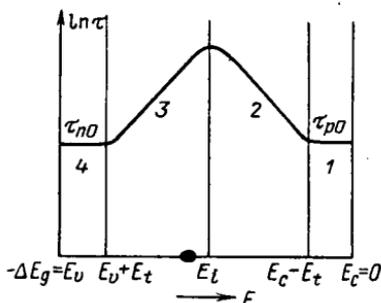


Рис. 6-7. Изменение времени жизни в зависимости от положения уровня Ферми.

Учет этих неравенств приводит, что

$$\tau = \tau_{p0} = \frac{1}{\gamma_p N_t}. \quad (6-85)$$

В сильно легированном полупроводнике *n*-типа величина $\tau = \tau_{p0}$ постоянна и определяется временем жизни дырок. В этом случае скорость процесса рекомбинации определяется числом актов захвата неравновесных дырок ловушками, занятыми электронами.

Вторая область. Полупроводник *n*-типа, но легирован слабо, так что уровень Ферми лежит ниже энергетического уровня ловушек, но выше середины запрещенной зоны, т. е. $E_i < F < (E_c - E_t)$. В этом случае имеют место следующие соотношения между концентрациями носителей:

$$n_0 \gg p_0; \quad n_0 \gg p_1; \quad n_0 < n_1.$$

Соответственно выражение для τ приобретает вид:

$$\tau = \tau_{p0} \frac{n_1}{n_0} = \tau_{p0} e^{\frac{E_c - E_t - F}{k_0 T}}. \quad (6-86)$$

Из соотношения (6-86) следует, что по мере понижения уровня Ферми величина времени жизни τ растет от значения τ_{p0} по экспоненциальному закону. При этом, чем ниже опускается уровень Ферми, тем меньше степень заполнения ловушек электронами, поэтому вероятность захвата дырки снижается, что ведет к увеличению времени жизни неравновесных электронно-дырочных пар.

Третья область. В слабо легированном полупроводнике *p*-типа положение уровня Ферми определяется неравенством

$$(E_v + E_t) < F < E_i,$$

что ведет к следующим соотношениям: $p_0 \gg n_0; p_0 \gg p_1; n_1 > p_0$. Отсюда получаем формулу для τ :

$$\tau = \tau_{p0} \frac{n_1}{p_0} = \tau_{p0} \frac{N_c}{N_v} e^{\frac{F - E_t + \Delta E_g}{k_0 T}}. \quad (6-87)$$

Отсюда видно, что по мере снижения уровня Ферми наблюдается экспоненциальное уменьшение времени жизни, что объясняется следующим образом. В слабо легированном полупроводнике *p*-типа почти все ловушки пусты и легко захватывают электроны из зоны проводимости. Но так как концентрация свободных дырок невелика часть электронов уйдет с ловушек обратно в зону проводимости. С понижением уровня Ферми увеличивается количество свободных дырок и растет вероятность их рекомбинации с электронами, попавшими на ловушки. При этом интенсивность процесса ионизации электронов снижается и время жизни уменьшается.

Четвертая область. Сильно легированный материал p -типа, для которого положение уровня Ферми определяется неравенством вида $E_v < F < (E_v + E_t)$. В этом случае

$$p_0 \gg n_0; \quad p_0 \gg n_1; \quad p_0 \gg p_1,$$

а из (6-84) следует:

$$\tau = \tau_{n0} = \frac{1}{\gamma_n N_t}. \quad (6-88)$$

Таким образом, в сильно легированном полупроводнике p -типа величина $\tau = \tau_{n0}$ постоянна. Она определяется временем жизни электронов, т. е. временем жизни 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, количество которых в валентной зоне очень велико. Процесс обратного переброса электронов в зону проводимости практически полностью отсутствует и не влияет на величину времени жизни.

Рассмотрим теперь случай большого уровня инжекции неравновесных носителей заряда, когда $\Delta n \gg n_0; \Delta n \gg p_0$. Из равенства (6-84) получаем:

$$\tau_\infty \approx \tau_{p0} + \tau_{n0} = \frac{\gamma_n + \gamma_p}{\gamma_n \gamma_p N_t}. \quad (6-89)$$

Как видно из найденного соотношения, при большом уровне инжекции время жизни для рекомбинации через центры захвата τ_∞ не зависит от концентраций электронов и дырок. Оно определяется лишь количеством и свойствами ловушек.

Проведем анализ формулы (6-88), по которой $\tau_{n0} = 1/\gamma_n N_t$. Это равенство выражает собой тот факт, что каждый электрон проводимости в материале p -типа, перемещаясь в зоне, обладает определенной вероятностью быть захваченным ловушкой. Для количественной оценки этого процесса используем так называемое эффективное сечение захвата σ_n . Согласно выражению (4-45) вероятность захвата носителей заряда пропорциональна поперечному сечению захвата этого центра и тепловой скорости движения носителей

$$\gamma_n = \sigma_n v_0. \quad (6-90)$$

Следовательно, время жизни электрона в акцепторном полупроводнике

$$\tau = \tau_{n0} = \frac{1}{\gamma_n N_t} = \frac{1}{\sigma_n v_0 N_t} \quad (6-91)$$

уменьшается при увеличении концентрации ловушек и температуры, приводящей к росту v_0 .

6-8. ТЕМПЕРАТУРНАЯ ЗАВИСИМОСТЬ ВРЕМЕНИ ЖИЗНИ НОСИТЕЛЕЙ ЗАРЯДА ПРИ РЕКОМБИНАЦИИ ЧЕРЕЗ ЛОКАЛЬНЫЕ УРОВНИ

Рассмотрим теперь, как меняется время жизни носителей заряда при рекомбинации через центры захвата с изменением температуры в полупроводнике заданного состава. Это рассмотрение проведем на примере примесного полупроводника электронного типа.

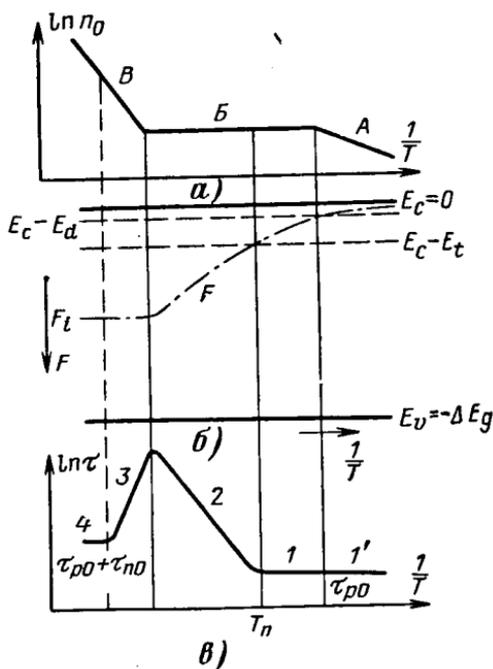


Рис. 6-8. Зависимость равновесной концентрации (а), уровня Ферми (б) и времени жизни (в) от температуры для донорного полупроводника.

В области А уровень Ферми, температурный ход которого изображен на рис. 6-8, б, лежит между уровнем донорной примеси и зоной 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 ловушек, концентрация электронов подчиняется неравенству $n_0 \gg n_1$ и соответственно $n_0 \gg p_0$ и $n_0 \gg p_1$. Следовательно, в этой области в соответствии с выражением (6-85) время жизни постоянно и равно $\tau = \tau_{p0}$ (участок 1' на рис. 6-8, в, на котором представлена температурная зависимость времени жизни).

По мере дальнейшего повышения температуры при переходе к области полной ионизации донорной примеси уровень Ферми понижается и достигает уровня ловушек. На этом участке концентрация электронов n_0 постоянна, а концентрация $n_1 =$

$= N_{ce} - \frac{E_t}{k_0 T}$ растет с температурой экспоненциально, но она по-преж-

нему значительно меньше равновесной концентрации n_0 . Кроме того, n_0 также больше p_0 и концентраций дырок p_1 и электронов n_1 , для которых положение уровня Ферми совпадает с уровнем ловушек. Согласно выражению (6-85) на этом участке время жизни также постоянно и равно τ_{p0} (участок 1, рис. 6-8, в).

При дальнейшем повышении температуры на участке 2 (рис. 6-8, в) уровень Ферми понижается, но находится ниже энергетического уровня ловушек. Здесь по-прежнему $n_0 = \text{const}$, а n_1 значительно увеличивается, поэтому из выражения (6-87) будем иметь:

$$\tau = \tau_{p0} \frac{n_1}{n_0} = \tau_{p0} \frac{N_c}{n_0} e^{-\frac{E_t}{k_0 T}} = \frac{\tau_{p0}}{n_0} 2 \left(\frac{2\pi m_n^* k_0}{h^2} \right)^{3/2} T^{3/2} e^{-\frac{E_t}{k_0 T}}, \quad (6-92)$$

т. е. время жизни увеличивается с ростом температуры. При этом оно увеличивается вплоть до температуры, при которой наступает собственная проводимость (участок 2 на рис. 6-8, в). Рост τ в этой области температур происходит потому, что с увеличением температуры имеет место интенсивная эмиссия электронов из ловушек. В этом случае уменьшается заполнение ловушек электронами, а следовательно, увеличивается время жизни неравновесных дырок, которое и определяет измеряемое время τ .

В области собственной проводимости с ростом температуры, хотя положение уровня Ферми практически остается постоянным (рис. 6-8, б), концентрация носителей в зонах резко увеличивается, поэтому увеличивается заполнение ловушек носителем, и время жизни уменьшается. В этой области

$$n_0 \approx p_0 \approx n_i = (N_c N_v)^{1/2} e^{-\frac{\Delta E_g}{2k_0 T}}. \quad (6-93)$$

В соответствии с этим соотношением (6-84) запишется в виде

$$\begin{aligned} \tau = \tau_{p0} \frac{n_i + n_1}{2n_i} + \tau_{n0} \frac{n_i + p_1}{2n_i} = \frac{\tau_{p0}}{2} \left(1 + \sqrt{\frac{N_c}{N_v}} e^{\frac{\Delta E_g}{2} - E_t} \right) + \\ + \frac{\tau_{n0}}{2} \left(1 + \sqrt{\frac{N_v}{N_c}} e^{-\frac{\Delta E_g}{2} + E_t} \right). \end{aligned} \quad (6-94)$$

Если уровень ловушек расположен достаточно далеко от середины запрещенной зоны, то в области начала собственной проводимости можно выделить участок, для которого выполняется неравенство $\left(\frac{\Delta E_g}{2} - E_t \right) \gg k_0 T$. Для этого случая уравнение (6-94) приобретает вид:

$$\tau = \frac{\tau_{p0}}{2} \left(\frac{N_c}{N_v} \right)^{1/2} e^{\frac{\Delta E_g}{2} - E_t}. \quad (6-95)$$

Соответствующий участок на кривой $\ln \tau (1/T)$ характеризуется уменьшением времени жизни (участок 3 на рис. 6-8, в).

При дальнейшем повышении температуры, когда в области собственной проводимости будет иметь место неравенство $(\Delta E_g/2 - E_t) \ll k_0 T$, уравнение (6-94) можно записать приближенно следующим образом:

$$\tau = \frac{\tau_{p0}}{2} \left(1 + \sqrt{\frac{N_c}{N_v}} \right) + \frac{\tau_{n0}}{2} \left(1 + \sqrt{\frac{N_v}{N_c}} \right). \quad (6-96)$$

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

$$\tau = \tau_{p0} + \tau_{n0} = \frac{\gamma_n + \gamma_p}{\gamma_n \gamma_p N_t}. \quad (6-97)$$

Таким образом, когда в области собственной проводимости с увеличением температуры концентрация свободных носителей становится очень высокой, время жизни определяется только параметрами центров захвата и не зависит от температуры (участок 4 на рис. 6-8, в).

Зная концентрацию электронов при насыщении, можно оценить величину E_t по температуре перехода T_n от области 1 к области 2 (рис. 6-8, в). При условии $T = T_n$, $F = E_c - E_t$ находим:

$$n_0 = N_c e^{-\frac{E_t}{k_0 T_n}} \quad (6-98)$$

или

$$E_t = k_0 T_n \ln \frac{N_c}{n_0}.$$

Экстраполируя прямую области 2 до пересечения с осью ординат, определяем в соответствии с равенством (6-92) значение $\ln(\tau_{p0} N_c/n_0)$ и, зная τ_{p0} из участка 1, находим $\ln N_c/n_0$ и подставляем это выражение в формулу (6-98). Следовательно, по снятой экспериментально зависимости $\ln \tau(1/T)$, подобной рис. 6-8, в, можно рассчитать глубину залегания уровня ловушек E_t .

6-9. ЦЕНТРЫ ПРИЛИПАНИЯ И ЦЕНТРЫ РЕКОМБИНАЦИИ

При анализе процессов рекомбинации носителей заряда мы учитывали как захваты электронов и дырок ловушками, так и тепловой заброс захваченных электронов и дырок с ловушек в зоны. При этом явление теплового возбуждения захваченных носителей 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. 6-9 изображена энергетическая схема полупроводника, содержащего локальные центры, характеризующиеся различными уровнями в запрещенной зоне. Если мы имеем состояние, отличное от термодинамического равновесия, значения концентраций сво-

бодных носителей определяются положением квазиуровней Ферми для электронов F_n и для дырок F_p . Если энергетический уровень ловушек расположен близко к дну зоны проводимости, вероятность тепловой ионизации велика. Такие центры находятся в состоянии непрерывного обмена электронами с зоной проводимости и не вносят существенного вклада в процессы рекомбинации. Подобные ловушки называют центрами прилипания для электронов, а соответствующие им уровни — уровнями прилипания. Аналогично вблизи верхнего края валентной зоны располагаются уровни прилипания для дырок.

Ловушки, которые осуществляют захват электронов и дырок, в результате чего происходит их рекомбинация, называются центрами рекомбинации, а соответствующие им уровни — уровнями рекомбинации. У таких ловушек обратный тепловой заброс маловероятен, поэтому они должны характеризоваться энергетическими уровнями, расположенными достаточно глубоко под дном зоны проводимости.

Чтобы охарактеризовать соотношение процессов прилипания и рекомбинации количественно, вводят коэффициент, выражающий собой вероятность захвата носителя противоположного знака.

Он равен отношению вероятности захвата к вероятности теплового заброса. Применительно к электронам зоны проводимости k_n будет равен отношению вероятности захвата отрицательно заряженной ловушкой дырки к вероятности теплового заброса электрона в зону проводимости. Как следует из выражения (6-70), скорость захвата ловушкой дырок определяется величиной $\gamma_p N_d f_t p$. Интенсивность процесса теплового возбуждения электронов с ловушек согласно формуле (6-69) равна $\gamma_n N_t f_t n_1$. Здесь γ_n , γ_p — вероятность захвата электрона и дырки соответственно; f_t — функция распределения для ловушки; p — концентрация дырок в валентной зоне; n_1 — концентрация электронов в зоне проводимости в том случае, когда уровень Ферми совпадает с уровнем ловушки.

Учитывая это, для коэффициента k_n получаем следующее выражение:

$$k_n = \frac{\gamma_p p}{\gamma_n n_1} = \frac{\gamma_p p}{\gamma_n N_c e^{-\frac{E_t}{k_0 T}}} \quad (6-99)$$

Ловушки, для которых вероятность рекомбинации больше вероятности теплового возбуждения, т. е. $k_n > 1$, будут являться

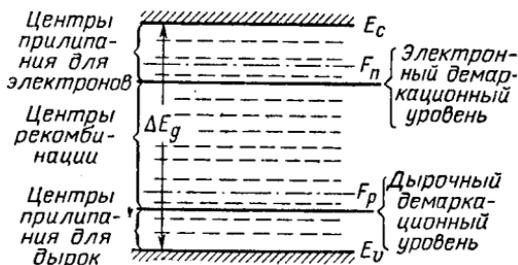


Рис. 6-9. Полупроводник, содержащий центры прилипания и центры рекомбинации.

центрами рекомбинации, а ловушки с $k_n < 1$ относятся к центрам прилипания. Энергетический уровень локального центра, который характеризуется значением $k_n = 1$, т. е. для которого вероятность рекомбинации и теплового заброса одинакова, называется электронным демаркационным уровнем E_{dn} . Его положение можно найти из условия

$$\gamma_p N_v e^{-\frac{\Delta E_g + F_p}{k_0 T}} = \gamma_n N_c e^{-\frac{E_{dn}}{k_0 T}}, \quad (6-100)$$

откуда имеем:

$$-E_{dn} = -\Delta E_g - F_p - k_0 T \ln \frac{\gamma_n N_c}{\gamma_p N_v}. \quad (6-101)$$

Из полученного соотношения следует, что положение электронного демаркационного уровня зависит от параметров ловушки (γ_n , γ_p , E_l), ширины запрещенной зоны, температуры и концентрации неосновных носителей заряда, захват которых завершает рекомбинацию. С увеличением уровня инжекции уменьшается величина $F_p - E_v$ и электронный демаркационный уровень приближается к дну зоны проводимости. В этом случае будет осуществляться переход центров прилипания в центры рекомбинации. С ростом же температуры демаркационный уровень сдвигается к середине запрещенной зоны и центры рекомбинации будут переходить в центры прилипания.

Для центров рекомбинации и центров прилипания, уровни которых расположены вблизи верхнего края валентной зоны, можно ввести коэффициент

$$k_p = \frac{\gamma_n n}{\gamma_p p_1} = \frac{\gamma_n n}{\gamma_p N_v e^{\frac{E_l - \Delta E_g}{k_0 T}}}. \quad (6-102)$$

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

$$\gamma_n N_c e^{\frac{F_n}{k_0 T}} = \gamma_p N_v e^{\frac{E_{dp} - \Delta E_g}{k_0 T}}$$

или

$$-E_{dp} = -\Delta E_g - F_n - k_0 T \ln \frac{\gamma_n N_c}{\gamma_p N_v}. \quad (6-103)$$

Таким образом, если энергетический уровень ловушек лежит выше электронного демаркационного уровня, они являются центрами прилипания для электронов. Ловушки, которые дают уровни ниже дырочного демаркационного уровня, служат центрами прилипания для дырок. Между ними располагаются энергетические уровни центров рекомбинации (рис. 6-9). Необходимо отметить, что такая классификация является весьма условной и схематичной. При ее проведении, в частности, считается, что отношение коэффициентов захвата γ_n и γ_p изменяется сравнительно слабо при перемещении

уровня ловушек. Центры прилипания могут также вносить некоторый вклад в процессы рекомбинации, который становится особенно заметным, если в полупроводнике отсутствуют локальные центры, дающие энергетические уровни в средней части запрещенной зоны.

Глава седьмая

ДИФФУЗИЯ И ДРЕЙФ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА

7-1. УРАВНЕНИЕ НЕПРЕРЫВНОСТИ

Рассмотрим поведение полупроводника, в котором в результате воздействия каких-либо внешних возбуждающих факторов возникли отступления от равновесных концентраций носителей заряда. В этом случае концентрации свободных электронов и дырок являются функциями пространственных координат и времени, т. е. $n(x, y, z, t)$ и $p(x, y, z, t)$, и скорости их изменения могут быть выражены с помощью уравнения непрерывности.

Вывод уравнения непрерывности проведем для одномерного однородного образца, в котором концентрация электронов меняется только в направлении оси x , как это представлено на рис. 7-1. Выделим в этом полупроводнике слой толщиной dx и сечением в 1 см^2 . Объем этого слоя будет численно равен dx . Пусть в момент времени t концентрация электронов была $n(x, t)$, тогда число электронов в этом объеме составит величину $n(x, t) dx$. В момент времени $t + dt$ их количество будет $n(x, t + dt) dx$ и изменение числа электронов за время dt в объеме dx составит:

$$n(x, t + dt) dx - n(x, t) dx = \frac{\partial n}{\partial t} dx dt. \quad (7-1)$$

Это изменение числа электронов может происходить в результате процессов генерации, рекомбинации, а также из-за наличия диффузии и дрейфа носителей заряда. Рассмотрим эти процессы.

1. Под процессами генерации понимаются все механизмы, посредством которых электроны, находящиеся в валентной зоне, на локальных уровнях примеси или на уровнях прилипания, могут быть возбуждены и переведены в зону проводимости. Во всех этих случаях нужно учитывать генерацию свободных носителей в результате тепловых процессов, а также внешних воздействий. К внеш-

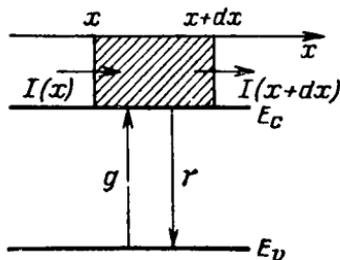


Рис. 7-1. К выводу уравнения непрерывности.

ним воздействиям относятся свет, ядерные частицы, неравновесное излучение, испускаемое самим полупроводником в его объеме, и т. п. Обозначим скорость этих двух типов генерации соответственно через g_0 и g . Тогда полная скорость генерации будет равна $g_0 + g$.

Будем считать, что в рассматриваемом случае возбуждение электронов происходит за счет света, и пусть число пар электрон — дырка, создаваемых светом за 1 сек в 1 см^3 , будет равно g . Тогда световая генерация в объеме dx за время dt создает электроны в количестве

$$g dx dt. \quad (7-2)$$

2. Изменение числа 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 рекомбинационных переходов при тепловом равновесии и r — скорость рекомбинации, вызванная внешними воздействиями. В этом случае, когда τ_n — время жизни неравновесных электронов, а их концентрация $\Delta n = n - n_0$, изменение числа электронов вследствие рекомбинации в объеме dx за время dt согласно выражению (6-18) составит величину

$$r dx dt = - \left(\frac{dn}{dt} \right)_r dx dt = \frac{n - n_0}{\tau_n} dx dt. \quad (7-3)$$

3. Изменение количества носителей заряда в объеме dx обусловлено, кроме того, процессами диффузии и дрейфа. Если $I_n(x, t)$ — число электронов, проходящих через 1 см^2 поверхности за 1 сек, то за время dt через границу слоя x в рассматриваемый объем войдет $I_n(x, t) dt$ электронов, а через границу $x + dx$ выйдет $I_n(x + dx, t) dt$ электронов. Следовательно, изменение числа электронов в объеме dx вследствие разности этих потоков будет равно:

$$I_n(x, t) dt - I_n(x + dx, t) dt = - \frac{\partial I_n}{\partial x} dx dt. \quad (7-4)$$

Полное изменение числа электронов составит величину

$$\frac{\partial n}{\partial t} dx dt = g dx dt - \frac{\partial I_n}{\partial x} dx dt - \frac{n - n_0}{\tau_n} dx dt, \quad (7-5)$$

откуда

$$\frac{\partial n}{\partial t} = g - \frac{\partial I_n}{\partial x} - \frac{n - n_0}{\tau_n}. \quad (7-6)$$

Уравнение (7-6) называют уравнением непрерывности. Аналогично уравнение непрерывности для дырок запишется в виде

$$\frac{\partial p}{\partial t} = g - \frac{\partial I_p}{\partial x} - \frac{p - p_0}{\tau_p}. \quad (7-7)$$

Потоки электронов I_n и дырок I_p можно выразить через плотности тока, которые с учетом знака носителя равны:

$$j_n = -eI_n \quad \text{и} \quad j_p = eI_p.$$

Когда концентрации носителей заряда являются функциями координат (x, y, z) , уравнения непрерывности запишутся через соответствующие токи следующим образом:

$$\frac{\partial n}{\partial t} = g + \frac{1}{e} \operatorname{div} \mathbf{j}_n - \frac{n - n_0}{\tau_n}; \quad (7-8)$$

$$\frac{\partial p}{\partial t} = g - \frac{1}{e} \operatorname{div} \mathbf{j}_p - \frac{p - p_0}{\tau_p}; \quad (7-9)$$

где $\operatorname{div} \mathbf{j} = \frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z}$.

В стационарных условиях концентрации электронов n и дырок p не меняются во времени, т. е. $\partial n / \partial t = \partial p / \partial t = 0$, уравнения непрерывности в одномерном случае принимают вид:

$$-\frac{1}{e} \frac{\partial j_n}{\partial x} = g - \frac{n - n_0}{\tau_n}; \quad (7-10)$$

$$\frac{1}{e} \frac{\partial j_p}{\partial x} = g - \frac{p - p_0}{\tau_p}. \quad (7-11)$$

Уравнения (7-10) и (7-11) выражают собой условия сохранения числа носителей заряда, т. е. в стационарном состоянии поток носителей заряда, вытекающих из объема, равен количеству носителей, созданных внешним возбуждением, за вычетом носителей прорекомбинировавших в этом объеме.

7-2. ДИФFUЗИОННЫЙ И ДРЕЙФОВЫЙ ТОКИ

Рассмотрим теперь неоднородный полупроводник, в котором концентрации электронов $n(\mathbf{r})$ и дырок $p(\mathbf{r})$ меняются 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.

Пусть в полупроводнике концентрация носителей заряда возрастает в направлении оси x , как это изображено на рис. 7-2. 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 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 придут носители заряда из слоя 2. Так как их число в слое 2 больше, чем в 1, то обратный поток будет больше прямого. Если $n(x - \frac{dx}{2})$ —

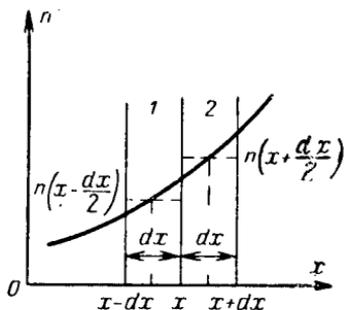


Рис. 7-2. К расчету градиента концентрации носителей заряда.

средняя концентрация электронов в слое 1, а $n\left(x + \frac{dx}{2}\right)$ — в слое 2, то разность концентраций в этих слоях будет равна:

$$n\left(x - \frac{dx}{2}\right) - n\left(x + \frac{dx}{2}\right) = -\frac{dn}{dx} dx. \quad (7-12)$$

Согласно уравнению (7-12) 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 уравнением вида

$$I_n = -D_n \frac{dn}{dx}, \quad (7-13)$$

где D_n — коэффициент диффузии электронов.

Аналогично диффузионный поток дырок

$$I_p = -D_p \frac{dp}{dx}, \quad (7-14)$$

где D_p — коэффициент диффузии дырок.

Потоки электронов и дырок, как следует из уравнений (7-13) и (7-14), текут в сторону меньших концентраций носителей заряда.

Диффузионным потокам носителей заряда соответствуют диффузионные токи электронов $j_{n \text{ дифф}}$ и дырок $j_{p \text{ дифф}}$:

$$j_{n \text{ дифф}} = eD_n \frac{dn}{dx}; \quad (7-15)$$

$$j_{p \text{ дифф}} = -eD_p \frac{dp}{dx}. \quad (7-16)$$

В том случае, если n является функцией координат (x, y, z) , диффузионный ток в векторной форме имеет вид для электронов

$$j_{n \text{ дифф}} = eD_n \text{grad } n(\mathbf{r}) \quad (7-17)$$

и для дырок

$$j_{p \text{ дифф}} = -eD_p \text{grad } p(\mathbf{r}). \quad (7-18)$$

Диффузионный ток, возникший из-за наличия градиента концентрации носителей, приведет к пространственному разделению зарядов. Разделение зарядов порождает статическое электрическое поле, которое создает дрейфовые токи электронов и дырок. В стационарном состоянии в каждой точке полупроводника дрейфовый ток будет уравновешивать диффузионный ток, поэтому суммарный ток при термодинамическом равновесии будет равен нулю.

Поместим неоднородный полупроводник во внешнее постоянное электрическое поле \mathcal{E} . Под действием внешнего поля электроны и дырки приобретут направленное движение, в результате чего появятся электронные и дырочные токи проводимости. Если внешнее электрическое поле несильное и не изменяет характера движения носителей заряда, то дрейфовые составляющие плотности тока

запишутся на основании закона Ома в виде:

$$\left. \begin{aligned} j_{n\text{др}} &= en \mu_n \mathcal{E}; \\ j_{p\text{др}} &= ep \mu_p \mathcal{E}. \end{aligned} \right\} \quad (7-19)$$

Полный ток будет складываться из диффузионного и дрейфового токов. Для электронов и дырок он будет равен:

$$j_n = j_{n\text{др}} + j_{n\text{диф}} = en \mu_n \mathcal{E} + eD_n \frac{dn}{dx}; \quad (7-20)$$

$$j_p = j_{p\text{др}} + j_{p\text{диф}} = ep \mu_p \mathcal{E} - eD_p \frac{dp}{dx}. \quad (7-21)$$

Таким образом, плотность общего тока j в любой точке неоднородного полупроводника в любой момент времени будет определяться уравнением

$$j = j_n + j_p = e(n\mu_n + p\mu_p) \mathcal{E} + e \left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx} \right). \quad (7-22)$$

Необходимо отметить, что диффузионный ток существует только в полупроводниках. Это происходит потому, что в полупроводниках концентрация электронов и дырок может меняться в широких пределах при постоянном значении суммарной концентрации зарядов. В металлах концентрация электронов практически постоянна.

7.3. СООТНОШЕНИЕ ЭЙНШТЕЙНА

В неоднородном полупроводнике при термодинамическом равновесии ток равен нулю, т. е. $j = j_n + j_p = 0$. В этом случае токи проводимости уравниваются диффузионные токи, и на основании (7-20) для электронов можно записать:

$$n \mu_n \mathcal{E}_{\text{ст}} = -D_n \frac{dn}{dx}. \quad (7-23)$$

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 вырождения концентрация электронов в зоне проводимости будет удовлетворять соотношению Больцмана вида

$$n = N_c e^{-\frac{E_c - F + U}{k_0 T}} = n_0 e^{\frac{e\varphi}{k_0 T}}, \quad (7-24)$$

где $n_0 = N_c e^{-\frac{E_c - F}{k_0 T}}$ — равновесная концентрация электронов; φ — электростатический потенциал.

Учитывая, что $\mathcal{E}_{\text{ст}} = -\frac{d\varphi}{dx}$, и подставляя значения n и dn/dx в уравнение (7-23), получаем:

$$- \mu_n n_0 e^{\frac{e\varphi}{k_0 T}} \frac{d\varphi}{dx} = -D_n \frac{e}{k_0 T} n_0 e^{\frac{e\varphi}{k_0 T}} \frac{d\varphi}{dx}, \quad (7-25)$$

откуда для электронов будем иметь:

$$\frac{\mu_n}{D_n} = \frac{e}{k_0 T}. \quad (7-26)$$

Аналогично для дырок

$$\frac{\mu_p}{D_p} = \frac{e}{k_0 T}. \quad (7-27)$$

Соотношение, связывающее коэффициент диффузии носителей заряда и их подвижность 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 за малое время, намного меньше их времени жизни, обмениваясь энергией с решеткой, приходят в тепловое равновесие с решеткой, и распределение их по энергии при отсутствии вырождения не отличается от распределения равновесных носителей заряда.

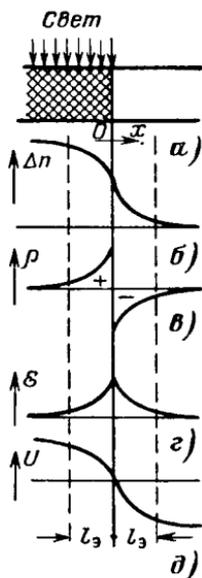


Рис. 7-3. Распределение концентрации (б), объемного заряда (в), электрического поля (г), потенциала (д) при частичном освещении донорного полупроводника.

Рассмотрим диффузию и дрейф неравновесных носителей заряда в случае монополярной проводимости, когда свободные носители возникают только в результате возбуждения их с уровней примеси. Пусть часть достаточно длинного однородного полупроводника, например, *n*-типа освещается слабопоглощающимся светом. У такого полупроводника в его освещенной части при $x < 0$ (рис. 7-3, а) имеет место равномерная генерация неравновесных электронов в результате перевода их светом с донорных уровней в зону проводимости. В этом случае концентрация электронов в зоне проводимости определяется концентрацией равновесных n_0 и неравновесных 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.

$$\Delta n = \Delta N_d^+. \quad (7-28)$$

Это равенство выражает собой условие электронейтральности в случае монополярной проводимости.

Неравновесные электроны из освещенной части образца будут диффундировать в неосвещенную часть (рис. 7-23, б). Вследствие

этого нарушается электронейтральность в некоторой области полупроводника и возникнет объемный заряд, а следовательно, и электрическое поле. В темной части образца, куда в результате диффузии пришли избыточные электроны, возникнет отрицательный объемный заряд ρ , а в освещенной области, откуда они ушли, — положительный заряд, обусловленный ионами донорной примеси (рис. 7-3, в). Эти заряды создадут электрическое поле $\mathcal{E}_{ст}$, направленное так, что оно будет препятствовать диффузии неравновесных электронов (рис. 7-3, г). Таким образом, возникновение диффузионного тока

$j_{диф} = eD_n \frac{dn}{dx}$ приводит к появлению статического электрического поля $\mathcal{E}_{ст}$, а следовательно, и тока проводимости $j_{др} = en\mu_n \mathcal{E}_{ст}$, направленного против тока диффузии. Поскольку полупроводник изолирован, то в стационарном состоянии полный ток равен нулю;

$$j = j_{диф} + j_{др} = en\mu_n \mathcal{E}_{ст} + eD_n \frac{dn}{dx} = 0. \quad (7-29)$$

Из равенства (7-29) можно определить напряженность статического электрического поля. Расчет величины $\mathcal{E}_{ст}$ проведем для случая малого уровня светового возбуждения, когда концентрация избыточных электронов мала по сравнению с равновесной, т. е. $\Delta n \ll n_0$. Используя соотношение Эйнштейна, будем иметь:

$$\mathcal{E}_{ст} = - \frac{D_n}{\mu_n n} \frac{d\Delta n}{dx} = - \frac{k_0 T}{e(n_0 + \Delta n)} \frac{d\Delta n}{dx} \approx - \frac{k_0 T}{en_0} \frac{d\Delta n}{dx}. \quad (7-30)$$

Продифференцировав (7-30), получим:

$$\frac{d\mathcal{E}_{ст}}{dx} = - \frac{k_0 T}{en_0} \frac{d^2 \Delta n}{dx^2}. \quad (7-31)$$

Величину $\frac{d\mathcal{E}_{ст}}{dx}$ можно найти, воспользовавшись уравнением Пуассона:

$$\frac{d\mathcal{E}_{ст}}{dx} = \frac{1}{\epsilon\epsilon_0} \rho = - \frac{e}{\epsilon\epsilon_0} \Delta n, \quad (7-32)$$

где $\rho = -e\Delta n$ — объемный заряд в неосвещенной части образца.

Из равенств (7-31) и (7-32) следует, что

$$\frac{d^2 \Delta n}{dx^2} - \frac{e^2 n_0}{\epsilon\epsilon_0 k_0 T} \Delta n = 0. \quad (7-33)$$

Если ввести обозначение

$$l_3 = \sqrt{\frac{\epsilon\epsilon_0 k_0 T}{e^2 n_0}}, \quad (7-34)$$

то уравнение (7-33) запишется следующим образом:

$$\frac{d^2 \Delta n}{dx^2} - \frac{\Delta n}{l_3^2} = 0. \quad (7-35)$$

Общее решение этого уравнения имеет вид:

$$\Delta n = C_1 e^{x/l_3} + C_2 e^{-x/l_3}, \quad (7-36)$$

где C_1 и C_2 — постоянные, определяемые из граничных условий; α_1 и α_2 — корни характеристического уравнения, равные

$$\alpha_1 = \frac{1}{l_s} \quad \text{и} \quad \alpha_2 = -\frac{1}{l_s}. \quad (7-37)$$

Для темной области полупроводника, в которой концентрация неравновесных электронов уменьшается по мере удаления от освещенной части образца, имеет смысл только член решения с отрицательным показателем степени, поэтому

$$\Delta n = \Delta n(0) e^{-\frac{x}{l_s}}. \quad (7-38)$$

Таким образом, в случае монополярной проводимости концентрация неравновесных носителей в темной части образца по мере удаления от освещенной области уменьшается по экспоненциальному закону с постоянной спада l_s , называемой радиусом (длиной) экранирования или дебаевским радиусом.

Длина экранирования, как следует из формулы (7-34), зависит от концентрации основных носителей, поэтому ее величина может меняться в широких пределах в зависимости от проводимости полупроводника. Например, для таких полупроводников, как германий и кремний, радиус экранирования составляет величину порядка $10^{-4} - 10^{-6}$ см. Следовательно, у этих полупроводников в случае монополярной проводимости неравновесные носители заряда диффундируют в глубь материала на очень малое расстояние из-за электростатических сил притяжения к неподвижным ионам, имеющим заряд противоположного знака.

Воспользуемся формулой (6-29), которая выражает изменение объемного заряда во времени $\rho = \rho_0 e^{-\frac{t}{\tau_\mu}}$. Если объемный заряд поделить на величину единичного заряда, то получим изменение концентрации неравновесных носителей во времени:

$$\Delta n = \Delta n(0) e^{-t/\tau_\mu}. \quad (7-38a)$$

Из сравнения формул (7-38) и (7-38a) следует, что распространение носителей в монополярном случае на расстояние длины экранирования l_s осуществляется в течение максвелловского времени релаксации τ_μ , которое в данном случае является эффективным временем установления диффузионно-дрейфового равновесия.

7.5. ДВИЖЕНИЕ НЕОСНОВНЫХ НОСИТЕЛЕЙ ЗАРЯДА

Рассмотрим более подробно вопрос о диффузии неосновных неравновесных носителей заряда при наличии внешнего однородного электрического поля \mathcal{E} .

Допустим, что имеется полупроводник p -типа ($p_0 \gg n_0$), поперечные размеры которого значительно меньше его длины. Часть полупроводника ($-l \leq x \leq 0$) освещается слабопоглощающим

светом, так что во всем объеме освещенной области образца происходит равномерная генерация электронов и дырок в результате ионизации основного вещества (рис. 7-4). Пусть концентрация основных носителей p_0 велика по сравнению с концентрацией неосновных неравновесных носителей Δn , время жизни которых τ_n . К образцу приложено однородное электрическое поле \mathcal{E} , так что носители могут дрейфовать в этом электрическом поле и диффундировать в неосвещенную область образца.

В освещенной области акцепторного полупроводника, где возбуждаются в равном количестве неравновесные электроны и дырки $\Delta n = \Delta p$, концентрация электронов, которые являются неосновными носителями, будет значительно больше, чем в неосвещенной части. Благодаря наличию значительного градиента концентрации электронов возникнет их диффузия. В неосвещенной части, куда продиффундируют электроны, появится объемный отрицательный заряд, электронейтральность образца нарушится. Этот объемный заряд создает электрическое поле, которое вызовет перераспределение носителей, так что в эту область в течение максвелловского времени релаксации будут подтянуты дырки и заряд избыточных электронов будет скомпенсирован.

Полная компенсация объемного заряда будет в том случае, если число добавочных дырок будет равным числу избыточных электронов. Таким образом, *избыточные неосновные носители — электроны, диффундируя в глубь полупроводника, вовлекают за собой равное количество основных носителей — дырок, так что объемный заряд не создается, а диффузия неосновных носителей электронов происходит как диффузия незаряженных частиц. По мере продвижения в глубь полупроводника избыточные электроны и дырки будут рекомбинировать и их концентрации с расстоянием будут убывать.*

Во внешнем электрическом поле избыточные неосновные носители — электроны будут дрейфовать, но их дрейф будет осуществляться таким образом, что будет сохраняться электронейтральность образца. При ускоряющем поле электроны будут затягивать за собой и дырки, а если поле будет замедлять электроны, то соответствующим образом произойдет и перераспределение основных носителей. Следовательно, и во внешнем электрическом поле избыточные неосновные носители дрейфуют как заряженные частицы, не создавая объемного заряда.

Концентрация избыточных электронов Δn может быть определена в результате решения уравнения непрерывности

$$\frac{\partial n}{\partial t} = g + \frac{1}{e} \frac{\partial j_n}{\partial x} - \frac{\Delta n}{\tau_n} \quad (7-39)$$

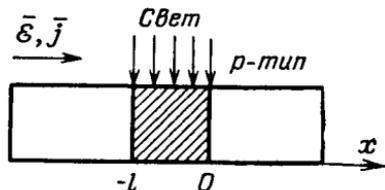


Рис. 7-4. Одномерная модель полупроводникового образца.

Найдем в стационарном случае ($\frac{\partial \Delta n}{\partial t} = 0$) распределение концентрации неосновных носителей вдоль оси x в неосвещенной части образца ($g = 0$). Для этого, продифференцировав выражение для плотности электронного тока (7-20) и подставив в уравнение (7-39), получим:

$$D_n \frac{d^2}{dx^2} \Delta n + \mu_n \mathcal{E} \frac{d}{dx} \Delta n - \frac{\Delta n}{\tau_n} = 0. \quad (7-40)$$

Это равенство справедливо для слабого внешнего электрического поля, когда D_n , μ_n и τ_n не зависят от поля.

Разделив обе части уравнения (7-40) на D_n и введя обозначения

$$L_D = \sqrt{D_n \tau_n} \quad (7-41)$$

и

$$L_{\mathcal{E}} = \tau_n \mu_n \mathcal{E}, \quad (7-42)$$

преобразуем уравнение (7-40) к виду

$$\frac{d^2}{dx^2} \Delta n + \frac{L_{\mathcal{E}}}{L_D^2} \frac{d}{dx} \Delta n - \frac{\Delta n}{L_D^2} = 0. \quad (7-43)$$

Общим решением этого уравнения будет:

$$\Delta n = C_1 e^{\alpha_1 x} + C_2 e^{\alpha_2 x}, \quad (7-44)$$

где C_1 и C_2 — постоянные, определяемые из граничных условий; α_1 , α_2 — корни характеристического уравнения;

$$\alpha^2 + \frac{L_{\mathcal{E}}}{L_D^2} \alpha - \frac{1}{L_D^2} = 0; \quad (7-45)$$

$$\alpha_{1,2} = \frac{-L_{\mathcal{E}} \pm \sqrt{L_{\mathcal{E}}^2 + 4L_D^2}}{2L_D^2}. \quad (7-46)$$

Учитывая уменьшение концентрации неосновных неравновесных носителей заряда по мере удаления от освещенной области образца, окончательно получаем:

при $x \geq 0$

$$\Delta n = C_1 e^{-\frac{x}{L_1}}, \quad (7-47)$$

где

$$L_1 = \frac{2L_D^2}{\sqrt{L_{\mathcal{E}}^2 + 4L_D^2} + L_{\mathcal{E}}}, \quad (7-48)$$

при $x \leq -l$

$$\Delta n = C_2 e^{\frac{x}{L_2}}, \quad (7-49)$$

где

$$L_2 = \frac{2L_D^2}{\sqrt{L_{\mathcal{E}}^2 + 4L_D^2} - L_{\mathcal{E}}}. \quad (7-50)$$

Таким образом, в области тени по обе стороны от освещенной части образца концентрация избыточных неосновных носителей снижается по экспоненциальному закону с постоянной спада L_1 и L_2 .

В отсутствие электрического поля, когда $L_{\mathcal{E}} = \mu_n \tau_n \mathcal{E} = 0$, имеет место только диффузия избыточных носителей заряда. Экспоненциальный спад концентрации избыточных носителей заряда, обусловленный рекомбинацией, будет симметричным по обе стороны от освещенной части образца и определяется величиной $L_D = \sqrt{D_n \tau_n}$, называемой диффузионной длиной неосновных носителей заряда (в данном случае электронов). L_D — это то среднее расстояние, на которое неравновесные носители при диффузии смещаются за время жизни. Следовательно, избыточная концентрация электронов будет изменяться с расстоянием по закону

$$\Delta n = \Delta n(0) e^{-\frac{x}{L_D}}. \quad (7-51)$$

Ток, который является только диффузионным током, будет уменьшаться по тому же закону, по которому изменяется избыточная концентрация электронов:

$$j = -e D_n \frac{dn}{dx} = \frac{e D_n}{L_D} \Delta n = \frac{e L_D}{\tau_n} \Delta n = j_D(0) e^{-x/L_D}. \quad (7-52)$$

Величина v_D , равная

$$v_D = \frac{L_D}{\tau_n} = \frac{L_D^2}{\tau_n L_D} = \frac{D_n}{L_D}, \quad (7-53)$$

называется диффузионной скоростью. Численно она равна скорости, с которой неравновесные носители проходят путь, равный диффузионной длине, за время их жизни.

Выражения (7-51)—(7-53) будут справедливы и при $\mathcal{E} \neq 0$, если, как следует из равенства (7-50), $L_{\mathcal{E}} \ll 2L_D$.

При наличии внешнего электрического поля $\mathcal{E} \neq 0$ постоянные спада L_1 и L_2 отличаются от диффузионной длины L_D . В зависимости от направления электрического поля они будут больше или меньше L_D .

В неосвещенной части, например, при $x \geq 0$

$$\begin{aligned} L_1 \Big|_{\mathcal{E} > 0} &= \frac{2L_D}{\sqrt{L_{\mathcal{E}}^2 + 4L_D^2} + L_{\mathcal{E}}} < L_D \Big|_{\mathcal{E} = 0} < \\ &< \frac{2L_D}{\sqrt{L_{\mathcal{E}}^2 + 4L_D^2} - L_{\mathcal{E}}} = L_2 \Big|_{\mathcal{E} < 0}. \end{aligned} \quad (7-54)$$

Величины L_1 и L_2 называются «диффузионной длиной вдоль поля» и «диффузионной длиной против поля», или длиной затягивания.

На рис. 7-5 изображено распределение неравновесных носителей заряда в отсутствие и при наличии внешнего электрического

поля. Поскольку при решении задачи предполагалось выполнение электронейтральности, по которому $\Delta n = \Delta p$, то изображенное на рис. 7-5 распределение носителей справедливо одновременно и для неравновесных электронов и для дырок.

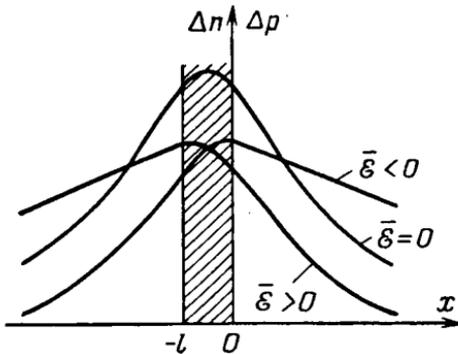


Рис. 7-5. Распределение концентрации неравновесных носителей заряда вдоль образца.

Рассмотрим теперь случай больших полей, для которых выполняется неравенство $L_{\mathcal{E}}^2 \gg 4L_D^2$. Это условие согласно (7-41), (7-42) и (7-53) можно записать следующим образом:

$$\frac{L_{\mathcal{E}}^2}{4L_D^2} = \frac{(\tau_n \mu_n \mathcal{E})^2}{4L_D^2} = \frac{\tau_n^2 v_d^3}{4L_D^2} = \frac{v_d^2}{4v_D^2} \gg 1. \quad (7-55)$$

Величина $L_{\mathcal{E}}$, численно равная пути, проходимому неравновесным электроном за время его жизни со скоростью дрейфа, называется длиной дрейфа неравновесных носителей заряда. Следовательно, поле \mathcal{E} будет большим, если длина дрейфа намного превосходит диффузионную длину.

Определим длины затягивания L_2 и L_1 для случаев, когда $L_{\mathcal{E}} > 0$ и $L_{\mathcal{E}} < 0$, т. е. для разного направления поля. Для случая $L_{\mathcal{E}} > 0$ из выражения (7-50) можно записать:

$$L_2 = \frac{2L_D^2}{L_{\mathcal{E}}} \frac{1}{\sqrt{1 + \frac{4L_D^2}{L_{\mathcal{E}}^2} - 1}} \approx \frac{2L_D^2}{L_{\mathcal{E}}} \frac{1}{\frac{2L_{\mathcal{E}}^2}{2L_{\mathcal{E}}^2}} = L_{\mathcal{E}}, \quad (7-56)$$

т. е. длина затягивания равна длине дрейфа:

$$L_2 = L_{\mathcal{E}} = \tau_n \mu_n \mathcal{E} = \tau_n v_d. \quad (7-57)$$

Изменение избыточной концентрации электронов будет определяться как

$$\Delta n = \Delta n(0) e^{-\frac{x}{L_{\mathcal{E}}}} = \Delta n(0) e^{-\frac{x}{\tau_n \mu_n \mathcal{E}}}. \quad (7-58)$$

Проведем анализ выражения (7-58).

1) Пусть в начале координат создается избыточная концентрация $\Delta n(0)$. При прохождении тока, вызванного сильным электрическим полем, когда $v_d \gg v_D$, избыточная концентрация носителей будет затягиваться полем в глубь полупроводника и полупроводник будет обогащаться неравновесными носителями заряда по сравнению со случаем $\mathcal{E} = 0$. Это явление называется инжекцией неравновесных носителей заряда. Инжекция электронов будет наблюдаться при $\mathcal{E} < 0$ (см. рис. 7-5), а инжекция дырок — при $\mathcal{E} > 0$.

2) Пусть в начале координат избыточная концентрация отрицательная, т. е. $n(0) - n_0 = \Delta n(0) < 0$. Согласно выражению (7-58) объем полупроводника будет обедняться носителями заряда. Это явление называется **экстракцией**. Экстракция электронов будет иметь место при $\mathcal{E} > 0$, а экстракция дырок — при $\mathcal{E} < 0$.

Определим теперь длину затягивания L_1 . Так как $L_{\mathcal{E}}^2 \gg 4L_D^2$, то из выражения (7-48) будем иметь:

$$L_1 = \frac{2L_D^2}{L_{\mathcal{E}}} \sqrt{1 + \frac{4L_D^2}{L_{\mathcal{E}}^2} + 1} \approx \frac{2L_D^2}{L_{\mathcal{E}}} \frac{1}{2 + \frac{2L_D^2}{L_{\mathcal{E}}^2}} \approx \frac{L_D^2}{L_{\mathcal{E}}}. \quad (7-59)$$

Из соотношения (7-59) следует, что с ростом поля (по модулю) длина затягивания стремится к нулю. При этом, если в начале координат $\Delta n(0) > 0$, то согласно (7-47) происходит уменьшение концентрации неравновесных носителей заряда. При $\Delta n(0) < 0$ имеет место увеличение концентрации носителей в объеме полупроводника. Эти явления соответственно называются **эксклюзией** и **аккумуляцией** неравновесных носителей заряда.

7-6. БИПОЛЯРНЫЙ КОЭФФИЦИЕНТ ДИФфуЗИИ И БИПОЛЯРНАЯ ДРЕЙФОВАЯ ПОДВИЖНОСТЬ

Рассмотрим теперь диффузию и дрейф неравновесных носителей в случае биполярной проводимости, когда неравновесные носители возникают в результате ионизации основного вещества. В этом случае $\Delta n = \Delta p$. Изучение диффузии и дрейфа проведем для однородного биполярного полупроводника, в котором имеется достаточно большое количество электронов и дырок, т. е. для полупроводника с проводимостью, близкой к собственной.

Пусть в таком одномерном полупроводнике существует неоднородное распределение носителей $n(x) = n_0 + \Delta n(x)$ 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 наличию градиента концентраций носителей возникнут диффузионные и дрейфовые токи электронов и дырок, которые будут описываться уравнениями (7-20) и (7-21), а полный ток будет равен:

$$j = j_n + j_p = e(n\mu_n + p\mu_p)\mathcal{E} + e\left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx}\right). \quad (7-60)$$

В изолированном полупроводнике в стационарном состоянии полный ток равен нулю ($j = 0$). При этом в каждой точке образца диффузионные токи уравновешиваются дрейфовыми токами, обусловленными статическим электрическим полем $\mathcal{E}_{ст}$. Величину $\mathcal{E}_{ст}$ найдем из уравнения (7-60), если положить $j = 0$:

$$\mathcal{E}_{ст} = \frac{D_p \frac{dp}{dx} - D_n \frac{dn}{dx}}{n\mu_n + p\mu_n} = \frac{D_p \frac{dp}{dx} - D_n \frac{dn}{dx}}{\mu_n(n_0 + \Delta n) + \mu_p(p_0 + \Delta p)} \quad (7-61)$$

Из (7-61) видно, что при отсутствии электрического тока ($j = 0$) возникающее в результате диффузии носителей статическое электрическое поле $\mathcal{E}_{ст}$ будет тем меньше, чем больше равновесные концентрации носителей n_0 и p_0 , поскольку дрейфовый ток создается и электронами и дырками. Если электропроводность полупроводника достаточно велика и концентраций неравновесных носителей заряда Δn и Δp значительно меньше концентрации равновесных носителей, т. е. $\Delta n \ll n_0$ и $\Delta p \ll p_0$, то в первом приближении можно считать, что 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 вовлекают с собой в процессе диффузии и носители противоположного знака в равной концентрации. В этом случае процесс диффузии неравновесных электронов и дырок происходит свободно, как диффузия неравновесных нейтральных пар электрон — дырка, характеризующихся одним временем жизни τ .

Пусть образец, в котором созданы неравновесные пары электрон — дырка, помещен во внешнее однородное электрическое поле \mathcal{E} , значительно превосходящее 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 плотности токов должны быть записаны как для электронов, так и для дырок

$$\frac{\partial \Delta n}{\partial t} = \frac{1}{e} \frac{\partial j_n}{\partial x} - \frac{\Delta n}{\tau}; \quad (7-62)$$

$$\frac{\partial \Delta p}{\partial t} = -\frac{1}{e} \frac{\partial j_p}{\partial x} - \frac{\Delta p}{\tau}; \quad (7-63)$$

$$\begin{aligned} j_n &= \sigma_n \mathcal{E} + e D_n \frac{dn}{dx}; \\ j_p &= \sigma_p \mathcal{E} - e D_p \frac{dp}{dx}, \end{aligned} \quad (7-64)$$

где $\sigma_n = en\mu_n$ и $\sigma_p = ep\mu_p$ — электронная и дырочная составляющие 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 электронейтральности 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 биполярной дрейфовой подвижностью и биполярным коэффициентом диффузии.

Для того чтобы определить $\mu_{\mathcal{E}}$ и D , запишем уравнения непрерывности (7-62) и (7-63), подставив в них значения токов j_n и j_p из

уравнения (7-64),

$$\frac{\partial \Delta n}{\partial t} = D_n \frac{d^2 \Delta n}{dx^2} + \mu_n \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau}; \quad (7-65)$$

$$\frac{\partial \Delta p}{\partial t} = D_p \frac{d^2 \Delta p}{dx^2} - \mu_p \mathcal{E} \frac{\partial \Delta p}{\partial x} - \frac{\Delta p}{\tau}. \quad (7-66)$$

Умножим уравнение (7-65) на σ_p , а соотношение (7-66) — на σ_n и сложим оба уравнения. В результате, учитывая, что $\Delta n = \Delta p$, получаем:

$$\frac{\partial \Delta n}{\partial t} = \frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} \frac{\partial^2 \Delta n}{\partial x^2} + \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau}. \quad (7-67)$$

Для стационарного случая, когда $\frac{\partial \Delta n}{\partial t} = 0$, уравнение (7-67) запишется в виде

$$\frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} \frac{\partial^2 \Delta n}{\partial x^2} + \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau} = 0. \quad (7-68)$$

Уравнение (7-68) по форме совпадает с уравнением (7-40) для диффузии и дрейфа неосновных носителей. Разница между ними лишь в том, что вместо коэффициента диффузии при второй производной и подвижности при первой производной в уравнении (7-68) стоят сложные величины, которые соответственно могут быть обозначены через D и $\mu_{\mathcal{E}}$.

Учитывая, что $n \approx n_0$ и $p \approx p_0$, а это справедливо при $\Delta n \ll n_0$ и $\Delta p \ll p_0$, и используя соотношение Эйнштейна для электронов и дырок $\left(\frac{\mu_n}{D_n} = \frac{\mu_p}{D_p} = \frac{e}{k_0 T} \right)$, коэффициент биполярной диффузии можно записать в виде

$$D = \frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} = \frac{n_0 + p_0}{\frac{n_0}{D_p} + \frac{p_0}{D_n}} = \frac{k_0 T}{e} \frac{n_0 + p_0}{\frac{n_0}{\mu_p} + \frac{p_0}{\mu_n}}, \quad (7-69)$$

а биполярную дрейфовую подвижность — следующим образом:

$$\mu_{\mathcal{E}} = \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} = \frac{p_0 - n_0}{\frac{n_0}{\mu_p} + \frac{p_0}{\mu_n}}. \quad (7-70)$$

Если воспользоваться соотношением Эйнштейна, то коэффициент биполярной диффузии D можно представить в виде

$$D = \frac{k_0 T}{e} \mu_D. \quad (7-71)$$

Из сопоставления с равенством (7-71) следует, что в уравнении (7-69) роль подвижности играет величина, равная

$$\mu_D = \frac{n_0 + p_0}{\frac{n_0}{\mu_p} + \frac{p_0}{\mu_n}} \quad (7-72)$$

и называемая биполярной диффузионной подвижностью.

Из уравнения (7-70) ясно, что величина $\mu_{\mathcal{E}}$ может изменять знак в зависимости от соотношения n_0 и p_0 и равна нулю при $n_0 = p_0$.

Для собственного полупроводника, у которого $n_0 = p_0$, имеем:

$$D = 2 \frac{D_n D_p}{D_n + D_p} = 2 \frac{k_0 T}{e} \frac{\mu_n \mu_p}{\mu_n + \mu_p}; \quad (7-73)$$

$$\mu_D = 2 \frac{\mu_n \mu_p}{\mu_n + \mu_p}; \quad (7-74)$$

$$\mu_{\mathcal{E}} = 0. \quad (7-75)$$

Из соотношений (7-73) и (7-75) следует, что у собственного полупроводника отсутствует дрейф носителей во внешнем электрическом поле, а диффузия их определяется коэффициентом, который зависит от коэффициента диффузии (или подвижностей) электронов и дырок. Отсутствие дрейфа свидетельствует о том, что внешнее электрическое поле в собственном полупроводнике не влияет на распределение концентрации носителей заряда.

Для примесных полупроводников, когда концентрация основных носителей заряда значительно превышает концентрацию неосновных носителей ($n_0 \gg p_0$ или $p_0 \gg n_0$), коэффициент биполярной диффузии D , биполярная диффузионная μ_D и биполярная дрейфовая $\mu_{\mathcal{E}}$ подвижности равны соответственно коэффициенту диффузии и подвижности неосновных носителей заряда. Например, для полупроводника n -типа ($n_0 \gg p_0$)

$$D = D_p,$$

а

$$\mu_D = |\mu_{\mathcal{E}}| = \mu_p. \quad (7-76)$$

В примесном полупроводнике диффузия и дрейф определяются неосновными носителями заряда.

Глава восьмая

КОНТАКТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

8-1. 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 мы проведем рассмотрение свойств электронного полупроводника, внесенного в однородное электрическое поле конденсатора (рис. 8-1).

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 появятся объемный заряд $\rho(\mathbf{r})$ и электрическое поле $\mathcal{E}(\mathbf{r})$. Изменение распределения концентрации носителей заряда, обуславливающее появление объемного заряда, будет иметь место в области, прилегающей к поверхности полупроводника. При таком подключении внешнего источника питания, как это изображено на рис. 8-1, в приповерхностной области полупроводника будет повышенная концентрация электронов (рис. 8-2, а), и, следовательно, возникнет отрицательный объемный заряд (рис. 8-2, б). Избыточная концентрация электронов, а значит, и объемный заряд будут уменьшаться с увеличением расстояния от поверхности в глубь полупроводника. Отрицательный объемный заряд порождает электрическое поле, напряженность которого \mathcal{E}_s будет максимальной на поверхности полупроводника (рис. 8-2, в). Электрическое поле изменит потенциальную энергию электрона на величину, равную $U(\mathbf{r}) = -e\varphi(\mathbf{r})$, где $\varphi(\mathbf{r})$ — потенциал поля (рис. 8-2, г). Следовательно, электрическое поле вызовет искривление зон энергии полупроводника таким образом, что

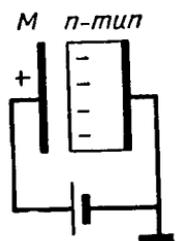


Рис. 8-1. Полупроводник в однородном электрическом поле.

$$\left. \begin{aligned} E_c(\mathbf{r}) &= E_c - U(\mathbf{r}); \\ E_v(\mathbf{r}) &= E_v - U(\mathbf{r}). \end{aligned} \right\} \quad (8-1)$$

При этом смещение испытывают все уровни энергии, в том числе и уровни примеси, лежащие в запрещенной зоне.

Поскольку полупроводник находится в состоянии термодинамического равновесия, то положение уровня Ферми постоянно (см. § 4-2), поэтому расстояние между уровнем Ферми и положением зон энергии изменится. Если это расстояние было без поля

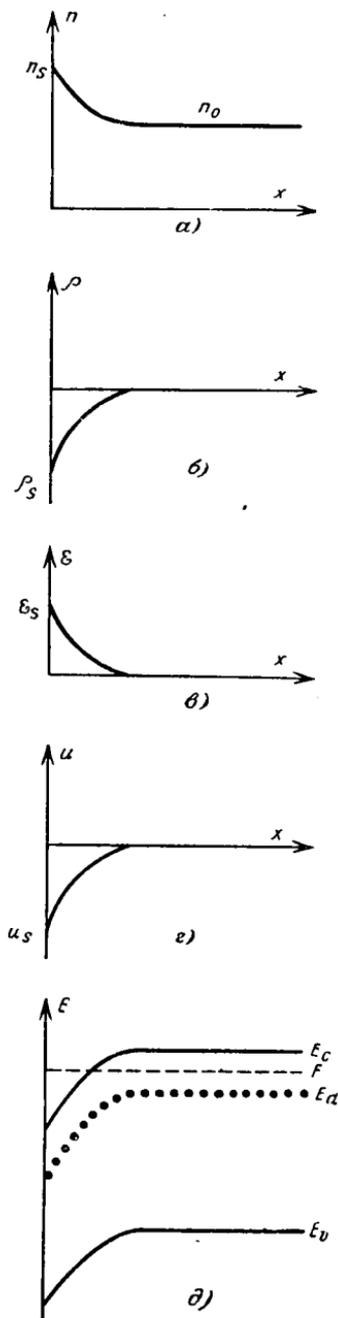
$$\mathcal{E}_c - F \text{ и } F - E_v, \quad (8-2)$$

то при наличии поля оно будет:

$$E_c - U(\mathbf{r}) - F \text{ и } F - [E_v - U(\mathbf{r})]. \quad (8-3)$$

Из сравнения (8-3) и (8-2) следует, если расстояние между E_c и F уменьшается на величину $U(\mathbf{r})$, то между F и E_v — увеличивается на ту же величину (рис. 8-2, д). Изменение расстояния между F и уровнем энергии должно привести к изменению распределения электронов по уровням. В рассматриваемом случае, как это изображено на рис. 8-2, д, если в глубине полупроводник электронный невырожденный и на уровнях донорной примеси имеются электроны (поскольку уровень Ферми лежит выше уровня примеси), то в приповерхностной области полупроводник также электронный, но вырожденный. Здесь уровень Ферми расположен в зоне проводимости.

При изменении направления приложенного напряжения к конденсатору соответственно изменится распределение концентрации



электронов, объемного заряда, электрического поля, потенциальной энергии электронов и расположения энергетических зон полупроводника. В данном случае, как это изображено на рис. 8-3, в приповерхностной области изменился тип проводимости, полупроводник из электронного превратился в дырочный. Слой полупроводника, в котором изменяется тип основных носителей заряда, называется инверсионным или обращенным слоем. В таком полупроводнике на некотором расстоянии от поверхности имеется слой с собственной проводимостью или *i*-слой и уровень Ферми лежит в середине запрещенной зоны. Около *i*-слоя располагается область, в которой меняется тип проводимости. Эта область носит название физического *p-n* перехода. Он исчезает при снятии внешнего электрического поля.

Расчет проведем для одномерного невырожденного полупроводника *n*-типа (рис. 8-3). Как известно, электрическое поле связано с объемным зарядом уравнением Пуассона

$$\frac{d}{dx} \epsilon \epsilon_0 \mathcal{E} = \rho(x). \quad (8-4)$$

Если поле выразить через потенциал $\mathcal{E} = -\frac{d\varphi}{dx}$, то уравнение Пуассона запишется в виде

$$\frac{d^2\varphi}{dx^2} = -\frac{1}{\epsilon \epsilon_0} \rho(x). \quad (8-5)$$

Пусть концентрация в глубине полупроводника ($x \rightarrow \infty$) равна n_0 , а на

Рис. 8-2. Изменение концентрации свободных электронов (а), объемного заряда (б), поля (в), потенциальной энергии электронов (г) и искривление зон донорного полупроводника (д) во внешнем электрическом поле.

поверхности n . Поскольку полупроводник невырожден, то

$$n = N_c e^{-\frac{E_c + U - F}{k_0 T}} = n_0 e^{-\frac{U}{k_0 T}}. \quad (8-6)$$

В приповерхностном слое объемный заряд определяется положительными ионами донорной примеси и свободными электронами. Будем считать, что донорная примесь полностью ионизирована, т. е. $N_d^+ = n_0$. Следовательно, объемный заряд в приповерхностном слое равен:

$$\rho = e(N_d^+ - n) = e(n_0 - n) = en_0 \left(1 - e^{-\frac{U}{k_0 T}}\right). \quad (8-7)$$

Ограничимся случаем малого искривления зон под действием поля, т. е. когда $|U| \ll k_0 T$. Разлагая выражение $e^{-\frac{U}{k_0 T}}$ в ряд и ограничиваясь первым членом разложения, на основании выражения (8-7) будем иметь:

$$\rho = -\frac{e^2 n_0}{k_0 T} \varphi. \quad (8-8)$$

Обозначим $\frac{e^2 n_0}{\epsilon \epsilon_0 k_0 T} = \frac{1}{l_s^2}$, тогда уравнение (8-8) примет вид:

$$\frac{d^2 \varphi}{dx^2} - \frac{1}{l_s^2} \varphi = 0. \quad (8-9)$$

Это уравнение имеет решение

$$\varphi = A_1 e^{-\frac{x}{l_s}} + A_2 e^{\frac{x}{l_s}}. \quad (8-10)$$

Так как в точке $x = 0$ $\varphi = -\varphi_s$, а при $x \rightarrow \infty$ $\varphi = 0$, значит, $A_2 = 0$ и $A_1 = -\varphi_s$, поэтому

$$\varphi = -\varphi_s e^{-\frac{x}{l_s}}. \quad (8-11)$$

Потенциальная энергия электрона

$$U(x) = -e\varphi(x) = U_s e^{-\frac{x}{l_s}}, \quad (8-12)$$

а объемный заряд на поверхности

$$\rho_s = \frac{en_0}{k_0 T} U_s. \quad (8-13)$$

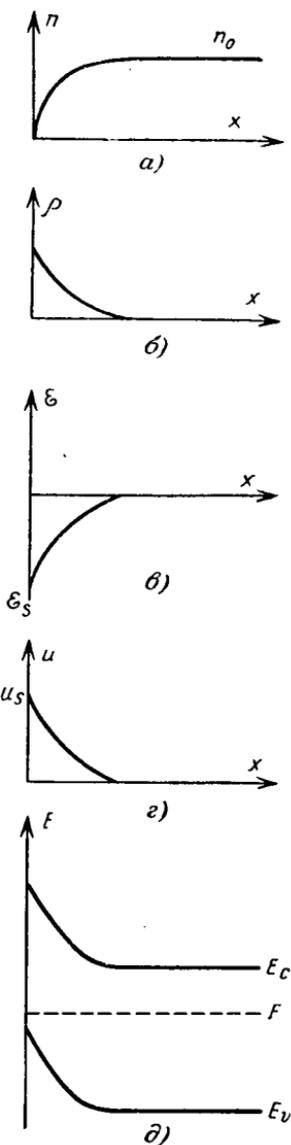


Рис. 8-3. Концентрация свободных электронов (а), объемный заряд (б), поле (в), потенциальная энергия электронов (г) и положение зон донорного полупроводника (д) при изменении направления внешнего электрического поля.

Из уравнений (8-12) и (8-13) видно, что при внесении полупроводника в электрическое поле в приповерхностной области его происходит искривление зон энергии и изменение концентрации электронов.

Знак объемной плотности заряда на поверхности полупроводника ρ_s зависит от знака U_s . При $U_s > 0$ зоны смещаются вверх и $\rho_s > 0$; при $U_s < 0$ зоны смещаются вниз и $\rho_s < 0$.

Дебаевская длина экранирования l_D — это то расстояние, на котором напряженность проникающего электрического поля в полупроводник уменьшается в e раз. Для металла, у которого $n_0 \approx 10^{22} \text{ см}^{-3}$ и $\epsilon = 1$, при комнатной температуре порядок длины экранирования составляет 10^{-8} см , т. е. под воздействием внешнего поля заряд сосредоточивается на поверхности металла. У чистого германия $n_0 = 10^{14} \text{ см}^{-3}$, а $\epsilon = 16$, электрическое поле проникает в глубь полупроводника на 4 мкм.

8-2. РАБОТА ВЫХОДА

Для перевода электрона из полупроводника в вакуум необходимо затратить энергию. Если за начало отсчета принять энергию E_a электрона, находящегося в вакууме и покоящегося относительно

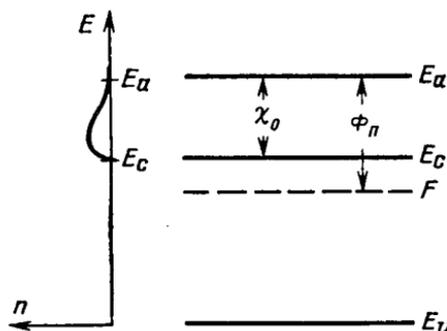


Рис. 8-4. Работа выхода электронов из полупроводника.

данного тела, то для перевода электрона с уровня E_c в вакуум без сообщения ему кинетической энергии требуется энергия χ_0 (рис. 8-4), равная

$$\chi_0 = E_a - E_c. \quad (8-14)$$

Эта энергия называется истинной работой выхода. Следовательно, на границе полупроводника существует энергетический барьер, препятствующий выходу электронов из кристалла.

Его могут покинуть лишь электроны, обладающие высокой энергией, достаточной для преодоления этого барьера. Очевидно, чем выше температура, тем больше будет электронов, способных преодолеть потенциальный барьер на границе полупроводника. Явление выхода из вещества электронов называется термоэлектронной эмиссией.

Подсчитаем количество электронов, способных преодолеть прямоугольный барьер высотой χ_0 и выйти из невырожденного полупроводника. Для этого кинетическая энергия электрона, имеющего скорость v_x , должна быть не меньше высоты потенциального барьера

$$\frac{m^* v_x^2}{2} \geq \chi_0. \quad (8-15)$$

Число квантовых состояний для кристалла единичного объема в интервале скоростей от \mathbf{v} до $\mathbf{v} + d\mathbf{v}$ определяется выражением (см. § 3-1)

$$dZ = 2 \left(\frac{m^*}{h} \right)^3 dv_x dv_y dv_z. \quad (8-16)$$

Поскольку для электронов, способных покинуть полупроводник, $E - F \gg k_0 T$, то их число будет:

$$dn = \int_0 dZ = 2 \left(\frac{m^*}{h} \right)^3 e^{-\frac{E-F}{k_0 T}} dv_x dv_y dv_z. \quad (8-17)$$

Будем считать, что все электроны, преодолевшие потенциальный барьер, не возвращаются в полупроводник. Если к поверхности кристалла движется поток со скоростью v_x , равной $v_x dn$, то плотность тока, текущего из вакуума в полупроводник, будет равна:

$$j = e \int_{v_{x \text{ мин}}}^{\infty} \int_{-\infty}^{\infty} v_x dn = 2e \left(\frac{m^*}{h} \right)^3 e^{\frac{F}{k_0 T}} \int_{v_{x \text{ мин}}}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{E}{k_0 T}} v_x dv_x dv_y dv_z. \quad (8-18)$$

С учетом, что полная энергия электрона $E = E_c + \frac{m^*}{2} (v_x^2 + v_y^2 + v_z^2)$, уравнение (8-18) запишется в виде

$$j = \frac{2e m^{*3} e^{-\frac{E_c - F}{k_0 T}}}{h^3} \int_{v_{x \text{ мин}}}^{\infty} e^{-\frac{m^* v_x^2}{2k_0 T}} v_x dv_x \int_{-\infty}^{\infty} e^{-\frac{m^* v_y^2}{2k_0 T}} dv_y \int_{-\infty}^{\infty} e^{-\frac{m^* v_z^2}{2k_0 T}} dv_z. \quad (8-19)$$

Для вычисления используем соотношение

$$\int_{-\infty}^{\infty} e^{-ay^2} dy = \sqrt{\frac{\pi}{a}}. \quad (8-20)$$

При интегрировании по v_x с учетом выражения (8-15) получим:

$$\int_{v_{x \text{ мин}}}^{\infty} e^{-\frac{m^* v_x^2}{2k_0 T}} v_x dv_x = \frac{k_0 T}{m^*} e^{-\frac{m^* v_{x \text{ мин}}^2}{2k_0 T}} = \frac{k_0 T}{m^*} e^{-\frac{\chi_0}{k_0 T}}. \quad (8-21)$$

Выражение (8-19) для плотности тока термоэлектронной эмиссии j на основании равенств (8-19) — (8-21) запишется в виде

$$j = \frac{4\pi e m^* k_0^3}{h^3} T^2 e^{-\frac{E_a - F}{k_0 T}} = A T^2 e^{-\frac{\Phi}{k_0 T}}, \quad (8-22)$$

где

$$A = \frac{4\pi e m^* k_0^3}{h^3},$$

$$\left\{ \Phi = \chi_0 + E_c - F = E_a - F. \right. \quad (8-23)$$

Величина Φ носит название термодинамической работы выхода электрона из полупроводника. Она составляет обычно несколько электрон-вольт.

Так как положение уровня Ферми в полупроводнике зависит от природы и концентрации примеси, и от температуры материала, то значение термодинамической работы выхода также будет зависеть от этих параметров. Выразим величину работы выхода для различных случаев, используя формулу (8-23) и соответствующие соотношения, определяющие положение уровня Ферми, которые были рассмотрены в § 3-6 и 3-7.

Для собственного полупроводника работа выхода Φ_i равна:

$$\Phi_i = \chi_0 + \frac{1}{2} \Delta E_g + \frac{k_0 T}{2} \ln \left(\frac{m_n^*}{m_p^*} \right)^{3/2}. \quad (8-24)$$

Из приведенного соотношения видно, что величина работы выхода собственного полупроводника определяется шириной запрещенной зоны, соотношением эффективных масс электронов, дырок и температурой.

Для донорного полупроводника в области слабой ионизации

$$\Phi_n = \chi_0 + \frac{E_c - E_d}{2} + \frac{k_0 T}{2} \ln \frac{\beta N_c}{N_d}, \quad (8-25)$$

а в области сильной ионизации примеси

$$\Phi_n = \chi_0 + k_0 T \ln \frac{N_c}{N_d}. \quad (8-26)$$

Здесь работа выхода уже не зависит от глубины залегания и степени вырождения донорного уровня, так как он лежит ниже уровня Ферми. Величина Φ_n определяется концентрацией примеси и температурой полупроводника.

Для акцепторного полупроводника в области слабой и сильной ионизации примеси имеем:

$$\Phi_p = \chi_0 + \Delta E_g - \frac{E_a - E_v}{2} - \frac{k_0 T}{2} \ln \left(\beta \frac{N_v}{N_a} \right); \quad (8-27)$$

$$\Phi_p = \chi_0 + \Delta E_g - k_0 T \ln \frac{N_v}{N_a}. \quad (8-28)$$

Из сравнения формул (8-25), (8-26), (8-27) и (8-28) следует, что работа выхода электронов из дырочного полупроводника больше, чем из электронного.

8-3. КОНТАКТ МЕТАЛЛ — МЕТАЛЛ. КОНТАКТНАЯ РАЗНОСТЬ ПОТЕНЦИАЛОВ

Пусть два куска различных металлов соединены одними концами вместе, а два других конца разделены вакуумным промежутком (рис. 8-5).

Рассмотрим вначале процессы, происходящие в контакте *C*. Предположим, что концентрация свободных электронов n_1 в первом металле больше, чем во втором n_2 , т. е. $n_1 > n_2$. В момент соприкосновения металлы начнут обмениваться электронами. Поскольку $n_1 > n_2$, то преимущественный переход электронов будет из первого металла во второй. На втором образце появится отрицательный заряд, а на первом — положительный, и между металлами возникнет разность потенциалов. В равновесном состоянии будет равенство потоков электронов, так как дрейфовый ток, вызванный полем разделенных зарядов, уравнивает диффузионный ток. Пусть при равновесном состоянии контактная разность равна U_i . Она называется внутренней контактной разностью.

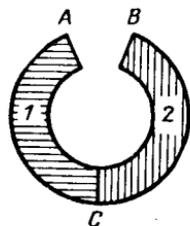


Рис. 8-5. Контакт двух разнородных металлов.

Потенциальная энергия электронов при переходе из первого образца во второй в точке *C* будет изменяться на величину eU_i (рис. 8-6, а).

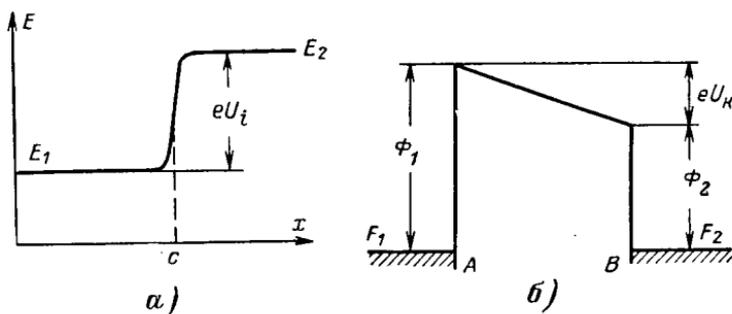


Рис. 8-6. Внутренняя (а) и внешняя (б) контактная разности потенциалов.

В соответствии с выражениями (8-17) и (8-18) вдоль оси x из первого металла во второй перейдут электроны, имеющие скорости v_{x1} , в количестве

$$dN_1 = 2 \frac{1}{m_1^* h^3} \rho_{x1} dp_{x1} \int_{-\infty}^{\infty} \int_e \frac{dp_{y1} dp_{z1}}{\frac{E_1 - F_1}{k_0 T} + 1}. \quad (8-29)$$

Количество перешедших электронов из второго в первый составит величину

$$dN_2 = 2 \frac{1}{m_2^* h^3} \rho_{x2} dp_{x2} \int_{-\infty}^{\infty} \int_e \frac{dp_{y2} dp_{z2}}{\frac{E_2 - F_2}{k_0 T} + 1}. \quad (8-30)$$

В условиях термодинамического равновесия число прямых переходов равно числу обратных переходов

$$dN_1 = dN_2. \quad (8-31)$$

Поскольку в этих условиях полные энергии электронов одинаковые, то

$$E_2 = E_1 + eU_1, \quad (8-32)$$

следовательно,

$$\frac{\partial E_1}{\partial p_{x1}} dp_{x1} = \frac{\partial E_2}{\partial p_{x2}} dp_{x2} \quad (8-33)$$

или

$$\frac{1}{m_1^*} p_{x1} dp_{x1} = \frac{1}{m_2^*} p_{x2} dp_{x2}. \quad (8-34)$$

Это означает, что двойные интегралы в уравнениях (8-29) и (8-30) равны между собой. Их равенство возможно лишь при условии, если

$$E_1 - F_1 = E_2 - F_2 = E_1 + eU_1 - F_2, \quad (8-35)$$

откуда следует:

$$eU_1 = F_2 - F_1. \quad (8-36)$$

Таким образом, *внутренняя контактная разность потенциалов определяется разностью энергий Ферми изолированных металлов.*

Для металлов уровень Ферми связан с концентрацией электронов выражением

$$F_0 = \frac{\hbar^2}{2m^*} \left(\frac{3n}{8\pi} \right)^{2/3}, \quad (8-37)$$

поэтому

$$eU_1 = \frac{\hbar^2}{2} \left(\frac{3}{8\pi} \right)^{2/3} \left(\frac{n_2^{2/3}}{m_2^*} - \frac{n_1^{2/3}}{m_1^*} \right). \quad (8-38)$$

Внутренняя контактная разность потенциалов определяется концентрацией электронов и их эффективными массами. Для металлов U_i невелико и составляет $10^{-3} - 10^{-2}$ в.

Определим теперь разность потенциалов, возникающую между металлами в вакуумном зазоре (на участке AB на рис. 8-5).

Предположим, что $\Phi_1 > \Phi_2$. В первый момент из второго металла ток термоэлектронной эмиссии будет больше, чем в обратном направлении, т. е. $j_2 > j_1$. В результате этого появится поле, препятствующее эмиссии электронов из второго образца. Направленный поток электронов будет до тех пор, пока система не придет в динамическое равновесие, характеризующееся постоянством уровня Ферми, и оба тока эмиссии не сравняются между собой:

$$j_{10} = j_{20} \quad (8-39)$$

или

$$AT^2 e^{-\frac{\Phi_1}{k_0 T}} = AT^2 e^{-\frac{\Phi_2 + eU_k}{k_0 T}}. \quad (8-40)$$

Внешняя контактная разность потенциалов, возникающая между металлами в вакуумном зазоре,

$$eU_k = \Phi_1 - \Phi_2 \quad (8-41)$$

равна разности работ выхода электронов из металлов. Это соотношение проиллюстрировано на рис. 8-6, б, где работа выхода изображена в виде соответствующих потенциальных барьеров на границе металла с вакуумом.

8-4. КОНТАКТ МЕТАЛЛ — ПОЛУПРОВОДНИК

Допустим, что имеем надежный контакт между металлом и электронным полупроводником. Если уровень Ферми металла F_m лежит ниже уровня Ферми полупроводника F_n , т. е. $\Phi_m > \Phi_n$, то поток электронов, вылетающих из полупроводника и попадающих в металл, в первый момент превышает поток электронов из металла. Металл заряжается отрицательно, а полупроводник — положитель-

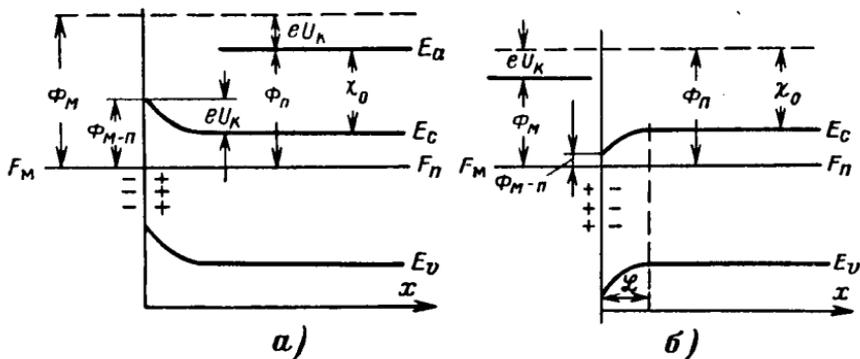


Рис. 8-7. Контакт металл — электронный полупроводник в случае (а) $\Phi_m > \Phi_n$ и (б) $\Phi_m < \Phi_n$.

но, в результате между контактирующими образцами возникает контактная разность потенциалов U_k и электрическое поле, препятствующее переходу электронов из полупроводника в металл. Направленный поток электронов будет иметь место до тех пор, пока уровни Ферми в системе не выровняются (рис. 8-7, а), после чего установится динамическое равновесие, характеризующееся равенством термоэлектронных токов

$$j_{m0} = j_{n0}, \quad (8-42)$$

откуда найдем, что

$$eU_k = \Phi_m - \Phi_n. \quad (8-43)$$

Контактная разность потенциалов практически полностью падает в приконтактной области полупроводника. Напряженность электрического поля в приповерхностном слое полупроводника, вызванного контактной разностью потенциалов, не превышает величины 10^6 в/см. Так как напряженность поля, обусловленного ионами кристаллической решетки, составляет приблизительно

10^8 в/см, небольшая добавка за счет контактного поля не может изменить ширину запрещенной зоны полупроводника. Однако в приповерхностном слое имеет место изгиб зон на величину eU_K . Если $\Phi_M > \Phi_n$, то зоны энергии в приконтактной области будут искривлены кверху (рис. 8-7, а и 8-8, а), а при $\Phi_n > \Phi_M$ — книзу (рис. 8-7, б и 8-8, б). Следовательно, если работа выхода электронов из металла больше, чем из полупроводника, то приконтактная область полупроводника обогащается дырками. При $\Phi_M < \Phi_n$ имеет место обогащение приконтактной области полупроводника электронами. Обогащение приконтактной области неосновными носителями заряда, как, например, в электронном полупроводнике при $\Phi_M > \Phi_n$ (рис. 8-7, а) или в дырочном при $\Phi_n < \Phi_M$ (рис. 8-8, б), сопровождается уменьшением проводимости. Слой с пониженной

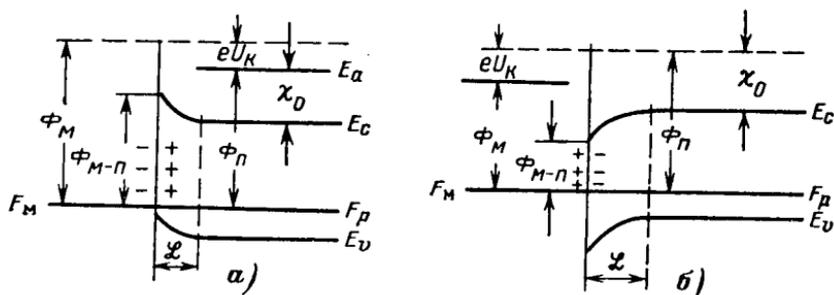


Рис. 8-8. Контакт металл — дырочный полупроводник в случае (а) $\Phi_M > \Phi_n$ и (б) $\Phi_M < \Phi_n$.

проводимостью называется з а п и р а ю щ и м. В том случае, если приконтактная область обогащается основными носителями, как, например, в электронном полупроводнике при $\Phi_n > \Phi_M$ (рис. 8-7, б) или в дырочном при $\Phi_M > \Phi_n$ (рис. 8-8, а), проводимость в приконтактной области повышается. Слой с повышенной проводимостью называется а н т и з а п и р а ю щ и м.

Выражение (8-43) можно преобразовать следующим образом:

$$eU_K = (\Phi_M - \chi_0) - (\Phi_n - \chi_0) = \Phi_{M-n} - \Phi_{n-M}. \quad (8-44)$$

Здесь $\Phi_{M-n} = \Phi_M - \chi_0$ — работа, которую необходимо совершить для перевода электрона из металла непосредственно в полупроводник; Φ_{n-M} — энергия, которую необходимо затратить для того, чтобы перевести электрон из полупроводника непосредственно в металл в первый момент после возникновения контакта (когда $U_K = 0$).

Предположим, что электрическое поле проникает в приконтактную область полупроводника на некоторую глубину \mathcal{L} . В приконтактной области энергия электронов на дне зоны проводимости равна $E_c + eU(x)$.

Согласно (8-7) объемный заряд в приконтактной области определяется соотношением

$$\rho = en_0 \left[1 - e^{-\frac{eU(x)}{k_0 T}} \right]. \quad (8-45)$$

Поскольку вся контактная разность потенциалов падает в приконтактной области полупроводника, можно считать, что $|eU(x)| \gg k_0 T$. Для этого случая выражение для плотности объемного заряда запишется в виде

$$\rho = en_0. \quad (8-45a)$$

Это означает, что на расстоянии \mathcal{L} , на которое проникает электрическое поле, из электронного полупроводника свободные электроны вытесняются полем и остается положительный заряд, определяемый концентрацией ионов донорной примеси.

Для области объемного заряда уравнение Пуассона с учетом уравнения (8-45a) примет вид:

$$\frac{d^2 U}{dx^2} - \frac{e}{\epsilon \epsilon_0} n_0 = 0. \quad (8-46)$$

Общим решением для этого уравнения является

$$U(x) = \frac{en_0}{2\epsilon\epsilon_0} (\mathcal{L} - x)^2 + A(\mathcal{L} - x) + B. \quad (8-47)$$

Так как поле проникает в полупроводник только на глубину \mathcal{L} , то уравнение (8-47) должно удовлетворять граничным условиям:

$$\begin{aligned} U(\mathcal{L}) &= 0; \\ \mathcal{E}(\mathcal{L}) &= - \left. \frac{dU}{dx} \right|_{x=\mathcal{L}} = 0. \end{aligned} \quad (8-48)$$

С учетом граничных условий (8-48) находим:

$$\begin{aligned} U(\mathcal{L}) = B &= 0; \\ \left. \frac{dU}{dx} \right|_{x=\mathcal{L}} &= -A = 0. \end{aligned} \quad (8-49)$$

Следовательно, в приконтактной области электростатический потенциал в зависимости от координаты x меняется следующим образом:

$$U(x) = \frac{en_0}{2\epsilon\epsilon_0} (\mathcal{L} - x)^2. \quad (8-50)$$

Для определения величины \mathcal{L} используем граничное условие в точке $x = 0$

$$U(0) = U_{\kappa} = \frac{1}{e} (\Phi_{\text{м}} - \Phi_{\text{п}}). \quad (8-51)$$

Это условие позволяет получить из уравнения (8-50) значение глубины проникновения поля:

$$\mathcal{L} = \sqrt{\frac{2\epsilon\epsilon_0 U_{\kappa}}{en_0}} = \sqrt{\frac{2\epsilon\epsilon_0 (\Phi_{\text{м}} - \Phi_{\text{п}})}{e^2 n_0}}. \quad (8-52)$$

Из полученной формулы следует, чем меньше степень легирования и чем больше разность работ выхода электронов из металла и полу-

проводника, тем больше глубина проникновения в полупроводник электрического поля, вызванного контактной разностью потенциалов. При разности работ выхода порядка 1 эВ глубина проникновения поля приблизительно в 10 раз превышает длину экранирования в случае контакта металл — вакуум — полупроводник.

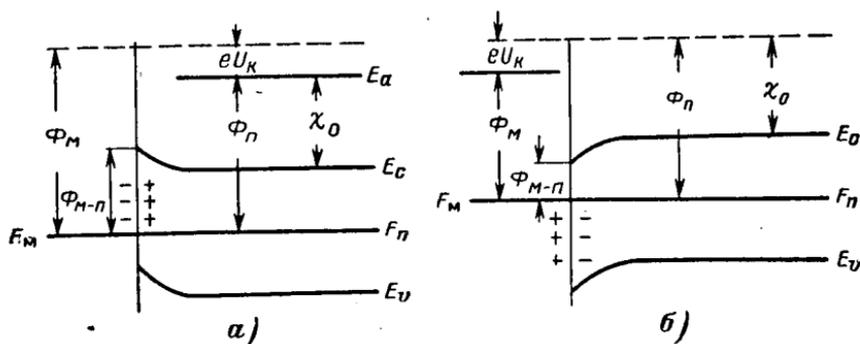


Рис. 8-9. Контакт металл — собственный полупроводник в случае (а) $\Phi_M > \Phi_n$ и (б) $\Phi_M < \Phi_n$.

При контакте металла с собственным полупроводником искривление зон сопровождается увеличением проводимости в приповерхностном слое. При этом, как видно из рис. 8-9, искривление зон различно для случая $\Phi_M > \Phi_n$ (рис. 8-9, а) и $\Phi_M < \Phi_n$ (рис. 8-9, б).

8-5. ВЫПРЯМЛЕНИЕ НА КОНТАКТЕ МЕТАЛЛ — ПОЛУПРОВОДНИК

Вернемся к рассмотренному выше случаю 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. 8-7, а). В глубине полупроводника ($x > \mathcal{L}$) концентрация свободных носителей

$$n_0 = N_c e^{-\frac{E_c - F}{k_0 T}} = N_c e^{\frac{\xi_0}{k_0 T}}, \quad (8-53)$$

где E_c — положение края зоны проводимости в глубине полупроводника; ξ_0 — положение уровня Ферми, отсчитываемого от нижнего края зоны проводимости в глубине полупроводника. Поскольку имеет место термодинамическое равновесие, то уровень Ферми для всей структуры один и тот же

$$F = E_c + \xi_0 = E_c + eU(x) + \xi(x). \quad (8-54)$$

При подключении источника напряжения через контакт металл — полупроводник будет протекать ток. Так как состояние полупроводника становится неравновесным, то для определения concentra-

ции электронов необходимо использовать квазиуровень Ферми (см. § 6-1)

$$F_n = E_c + eU(x) + \xi_1(x), \quad (8-55)$$

где $\xi_1(x)$ — положение квазиуровня Ферми, отсчитанного от дна зоны проводимости.

Концентрация электронов в области $0 < x < \mathcal{L}$ по аналогии с (8-53) может быть представлена в виде

$$n(x) = N_c e^{\frac{\xi_1(x)}{k_0 T}}. \quad (8-56)$$

Квазиуровень Ферми не постоянен для всего полупроводника. Чтобы определить как он изменяется, вычислим ток, протекающий через полупроводник. Плотность тока с учетом дрейфовой и диффузионной составляющих равна:

$$j = en\mu_n \mathcal{E} + eD_n \frac{dn}{dx}. \quad (8-57)$$

Выражая напряженность поля через потенциал и используя для величины n формулу (8-56), получаем:

$$j = n\mu_n \frac{d(eU)}{dx} + \frac{eD_n n}{k_0 T} \frac{d\xi_1}{dx}. \quad (8-58)$$

С помощью соотношения Эйнштейна выражение (8-58) преобразуется к виду

$$j = n\mu_n \frac{d(eU + \xi_1)}{dx}. \quad (8-59)$$

Сравнивая равенства (8-59) и (8-55), можно написать:

$$j = n\mu_n \frac{dF_n}{dx}. \quad (8-60)$$

Найдем, как изменится квазиуровень Ферми между двумя точками 1 и 2 полупроводника, в которых плотность тока одинакова:

$$\int_1^2 dF_n = \int_1^2 \frac{j dx}{n\mu_n} = ejs \int_1^2 \frac{dx}{en\mu_n s} = eI \int_1^2 dR. \quad (8-61)$$

Здесь $dR = dx/\sigma s$ — сопротивление участка полупроводника длиной dx с поперечным сечением s и удельной проводимостью $\sigma = en\mu_n$. Интегрируя уравнение (8-61), получаем:

$$\Delta F_n = F_{n2} - F_{n1} = eI (R_2 - R_1) = e(V_2 - V_1). \quad (8-62)$$

Таким образом, при наличии внешнего источника напряжения изменение квазиуровня Ферми определяется величиной приложенной разности потенциалов.

Определим теперь плотность тока, текущего через контакт металл — полупроводник, при различной полярности внешнего напряжения. Пусть к полупроводнику подключен отрицательный

полюс батареи, а к металлу — положительный (прямое подключение батареи). Будем полагать, что все внешнее напряжение падает в приповерхностном слое полупроводника, обладающем высоким сопротивлением. Положение уровня Ферми в глубине полупроводника (относительно уровня E_c) и в металле не изменяется, а в приконтактной области полупроводника он смещается на величину eV , где V — напряжение источника питания. Контактная разность потенциалов между металлом и полупроводником снижается и становится равной $e(U_k - V)$ (рис. 8-10, а).

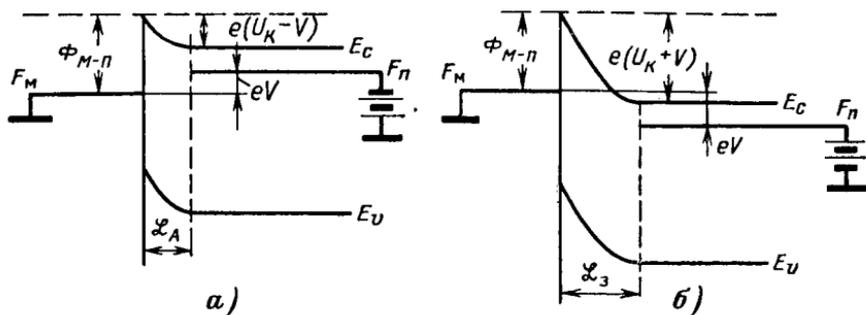


Рис. 8-10. Прямое (а) и обратное (б) подключение батареи к контакту металл — полупроводник.

В результате снижения потенциального барьера со стороны полупроводника равновесие в системе нарушится и потечет ток, направленный из металла в полупроводник

$$j = j_1 - j_2. \quad (8-63)$$

Со стороны металла высота потенциального барьера не изменилась, поэтому поток электронов из металла будет тот же, что и в равновесном состоянии, а поэтому

$$j_2 = AT^2 e^{-\frac{\phi_M}{k_0 T}}. \quad (8-64)$$

Поскольку со стороны полупроводника высота барьера понизилась и стала равной $e(U_k - V)$, то ток термоэлектронной эмиссии j_1 будет определяться выражением

$$j_1 = AT^2 e^{-\frac{\phi_n + eU_k - eV}{k_0 T}} = AT^2 e^{-\frac{\phi_M - eV}{k_0 T}}. \quad (8-65)$$

Результирующая плотность тока, текущего через контакт, будет равна:

$$j = AT^2 e^{-\frac{\phi_M}{k_0 T}} \left(e^{\frac{eV}{k_0 T}} - 1 \right) = j_s \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-66)$$

Величина $j_s = AT^2 e^{-\frac{\phi_M}{k_0 T}}$ носит название тока насыщения.

В случае обратного подключения батареи (рис. 8-10, б) со стороны полупроводника потенциальный барьер увеличится на величину приложенной разности потенциалов и выражение для тока, текущего через контакт, примет вид:

$$j = AT^2 e^{-\frac{\Phi_{\text{п}} + eU_{\text{к}} + eV}{k_0 T}} - AT^2 e^{-\frac{\Phi_{\text{м}}}{k_0 T}} = j_s \left(e^{-\frac{eV}{k_0 T}} - 1 \right). \quad (8-67)$$

Если при прямом подключении батареи величину V считать положительной, а при обратном подключении отрицательной, равенства (8-66) и (8-67) можно записать одним выражением

$$j = j_s \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-68)$$

Из (8-68) следует, что при прямом напряжении ($V > 0$) ток через контакт в зависимости от напряжения растет по экспоненциальному закону, а при обратном напряжении ($V < 0$) ток стремится к величине j_s . Таким образом, в рассмотренном случае контакт — металл — полупроводник обладает выпрямляющими свойствами.

Область полупроводника вблизи контакта представляет собой слой объемного заряда, толщина которого при отсутствии внешнего напряжения выражается формулой (8-52). В случае подключения батареи толщина слоя объемного заряда будет равна:

$$\mathcal{L} = \sqrt{\frac{2\epsilon\epsilon_0(U_{\text{к}} - V)}{\epsilon n_0}}, \quad (8-69)$$

где, как и в равенстве (8-68), знак V зависит от полярности подключенной батареи. При прямом смещении ($V > 0$) толщина слоя объемного заряда уменьшается по сравнению с толщиной при отсутствии батареи, а в случае обратного смещения ($V < 0$) увеличивается.

Одновременно с изменением толщины слоя объемного заряда \mathcal{L} в зависимости от приложенного напряжения изменяется и величина объемного заряда. Таким образом, контакт металл — полупроводник ведет себя подобно плоскому конденсатору, обладающему емкостью на единицу площади

$$C = \frac{\epsilon\epsilon_0}{\mathcal{L}} = \sqrt{\frac{\epsilon\epsilon_0\epsilon n_0}{2(U_{\text{к}} - V)}}. \quad (8-70)$$

8-6. ДИОДНАЯ ТЕОРИЯ ВЫПРЯМЛЕНИЯ

Для расчета вольт-амперной характеристики контакта металл — полупроводник используются две теории — д и о д н а я и д и ф ф у з и о н н а я. Если длина свободного пробега электронов l значительно меньше толщины запирающего слоя \mathcal{L} , носители заряда, проходя через этот слой, многократно рассеиваются. В этом случае необходимо учитывать как дрейфовую, так и диффузионную составляющие тока. Использование закона Ома для записи дрейфового тока возможно в том случае, если электрическое поле слабое, ус-

ловием чего является

$$\mathcal{E}_s \ll \frac{k_0 T}{e l}. \quad (8-71)$$

Максимальная напряженность поля будет на границе полупроводника с металлом ($x = 0$). Она с учетом равенств (8-50) и (8-52) равна:

$$\mathcal{E}_s = \mathcal{E}_{\text{макс}} = \left| -\frac{dU}{dx} \right|_{x=0} = \frac{en_0 \mathcal{L}}{\epsilon \epsilon_0} = \frac{2U_K}{\mathcal{L}}. \quad (8-72)$$

Подставляя найденное значение \mathcal{E}_s в неравенство (8-71), находим условие применимости диффузионной теории выпрямления:

$$\mathcal{L} \gg \frac{2eU_K}{k_0 T} l. \quad (8-73)$$

Если соотношение длины свободного пробега электронов l и толщины запирающего слоя \mathcal{L} таково, что выполняется обратное неравенство

$$l \gg \frac{k_0 T}{2eU_K} \mathcal{L}, \quad (8-74)$$

можно считать, что носители заряда пролетают эту запирающую область, почти не испытывая столкновений. При этом запирающий слой подобен вакуумному промежутку между электродами электронной лампы. В связи с этим теория, используемая в этом случае для расчета вольт-амперной характеристики контакта, называется **о д н о й**.

При рассмотрении контакта металл — полупроводник (§ 8-4) мы считали, что из металла способны выйти лишь электроны с энергией, превышающей величину 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 за счет туннельного эффекта.

Как следует из квантовой механики; прозрачность D потенциального барьера высотой Φ_0 и толщиной δ выражается формулой

$$D = e^{-\frac{2\delta}{\hbar} [2m^* (\Phi_0 - E_x)]^{1/2}}. \quad (8-75)$$

Здесь $E_x = \frac{m^* v_x^2}{2}$, где v_x — скорость в направлении к поверхности образца. Если принять, что $\Phi_0 \approx 1 \text{ эв}$; $\delta = 10^{-8} \text{ см}$, то при условии $E_x \ll \Phi_0$ получим $D = 1/3$. Это означает, что столь тонкие барьеры практически прозрачны для туннельного эффекта. Поэтому в дальнейшем мы 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 $e(U_K - V)$. Ток, обусловленный такими элек-

тронами, согласно выражению (8-22) будет выражаться соотношением:

$$j_1 = e \int_{v_{\text{мин}}}^{\infty} \int_{-\infty}^{\infty} v_x dn = \frac{4\pi m^* (k_0 T)^2}{h^3} e^{\frac{F-E_c}{k_0 T}} e^{-\frac{e(U_K - V)}{k_0 T}}. \quad (8-76)$$

Подставляя сюда выражения для равновесной концентрации электронов в глубине полупроводника

$$n_0 = 2 \left(\frac{2\pi m^* k_0 T}{h^2} \right)^{3/2} e^{\frac{F-E_c}{k_0 T}} \quad (8-77)$$

и их средней тепловой скорости

$$v_0 = \sqrt{\frac{8k_0 T}{\pi m^*}}, \quad (8-78)$$

получаем:

$$j_1 = \frac{1}{4} en_0 v_0 e^{-\frac{e(U_K - V)}{k_0 T}} = \frac{1}{4} en_s v_0 e^{\frac{eV}{k_0 T}}. \quad (8-79)$$

Здесь

$$n_s = n_0 e^{-\frac{eU_K}{k_0 T}} \quad (8-80)$$

— концентрация электронов на поверхности полупроводника при $V = 0$.

Со стороны металла величина потенциального барьера остается одна и та же (см. рис. 8-10), поэтому поток электронов из металла в полупроводник не зависит от величины внешнего смещения и равен потоку из полупроводника в состоянии равновесия:

$$j_2 = j_1 \Big|_{V=0} = \frac{1}{4} en_s v_0. \quad (8-81)$$

Суммарный ток через контакт будет определяться выражением

$$j = j_1 - j_2 = \frac{1}{4} en_s v_0 \left(e^{\frac{eV}{k_0 T}} - 1 \right) = j_s \left(e^{\frac{eV}{k_0 T}} - 1 \right), \quad (8-82)$$

где

$$j_s = \frac{1}{4} en_s v_0 \quad (8-83)$$

— ток насыщения для диодной теории выпрямления.

Как указывалось выше, диодная теория применима в том случае, когда выполняется условие (8-74). Толщина запирающего слоя \mathcal{L} тем меньше, чем больше концентрация свободных носителей n_0 , а длина свободного пробега электронов пропорциональна их подвижности

$$\mu = \frac{4el}{3\sqrt{2\pi m^* k_0 T}}. \quad (8-84)$$

Таким образом, диодную теорию следует применять для полупроводников с большой концентрацией носителей заряда, обладающих высокой подвижностью.

Например, для германия электронного типа: $\mu_n = 3\,900 \text{ см}^2/\text{в} \cdot \text{сек} = 0,39 \text{ м}^2/\text{в} \cdot \text{сек}$, $\epsilon = 16,5$ и $m^* = 0,3 m$. При $n_0 = 10^{14} \text{ см}^{-3}$ $eU_k = 0,3 \text{ эв}$, $\mathcal{L} = 2,3 \cdot 10^{-4} \text{ см}$, а $l = 1,5 \cdot 10^{-5}$ и отношение $\frac{k_0 T \mathcal{L}}{2eU_k l} = 0,6 < 1$. Если же повысить концентрацию электронов до $n_0 = 10^{16} \text{ см}^{-3}$, то $\frac{k_0 T \mathcal{L}}{2eU_k l} \approx 0,06$. В этом случае применение диодной теории выпрямления полностью оправдано.

В полупроводниках с малыми значениями концентрации и подвижности свободных носителей дело обстоит по-иному. В закиси меди p -типа: $\mu_p = 60 \text{ см}^2/\text{в} \cdot \text{сек}$, $\epsilon = 8,75$. При $p_0 = 10^{14} \text{ см}^{-3}$ $eU_k = 0,5 \text{ эв}$, следовательно, $\mathcal{L} = 2,2 \cdot 10^{-4} \text{ см}$ и $l = 4 \cdot 10^{-7} \text{ см}$. При этом хорошо выполняется неравенство

$$\frac{k_0 T \mathcal{L}}{2eU_k l} = 15 \gg 1.$$

В данном случае для расчета вольт-амперной характеристики контакта металл — полупроводник должна применяться диффузионная теория выпрямления.

8-7. ДИФФУЗИОННАЯ ТЕОРИЯ ВЫПРЯМЛЕНИЯ

В диффузионной теории выпрямления при определении тока, текущего через контакт металл — полупроводник, учитывается как дрейфовая, так и диффузионная составляющие тока. В этом случае для расчета вольт-амперной характеристики поступим следующим образом. Умножим обе части уравнения (8-57) на величину $e \frac{eU(x)}{k_0 T}$. Используя соотношение Эйнштейна и выражая напряженность поля через потенциал, будем иметь:

$$j e \frac{eU(x)}{k_0 T} = n \mu_n e \frac{eU(x)}{k_0 T} \frac{d(eU)}{dx} + \mu_n k_0 T e \frac{eU(x)}{k_0 T} \frac{dn}{dx} = \mu_n k_0 T \frac{d}{dx} \left[n e \frac{eU(x)}{k_0 T} \right]. \quad (8-85)$$

Принтегрируем обе части равенства (8-85) по x , величина которого может меняться от 0 до \mathcal{L} :

$$j \int_0^{\mathcal{L}} e \frac{eU(x)}{k_0 T} dx = \mu_n k_0 T \int_0^{\mathcal{L}} \frac{d}{dx} \left[n e \frac{eU(x)}{k_0 T} \right] dx. \quad (8-86)$$

Выразим отсюда величину j , которая не зависит от x ,

$$j = \frac{\mu_n k_0 T \left[n(\mathcal{L}) e \frac{eU(\mathcal{L})}{k_0 T} - n(0) e \frac{eU(0)}{k_0 T} \right]}{\int_0^{\mathcal{L}} e \frac{eU(x)}{k_0 T} dx}. \quad (8-87)$$

Для вычисления числителя полученной дроби обратимся к граничным условиям. В глубине полупроводника поле отсутствует, поэтому

$$U(\mathcal{L})=0 \text{ и } n(\mathcal{L})=n_0. \quad (8-88)$$

На границе раздела с металлом величина потенциала равна:

$$U(0)=U_K - V. \quad (8-89)$$

Рассмотрим случай, когда величина внешнего смещения сравнительно невелика и в приконтактной области всегда имеется запирающий слой:

$$U_K - V \gg \frac{k_0 T}{e}. \quad (8-90)$$

Можно показать, что при этом концентрация электронов на поверхности полупроводника практически не зависит от величины смещения, так что можно без большой ошибки положить:

$$n(0) = n_s = n_0 e^{-\frac{eU_K}{k_0 T}}. \quad (8-91)$$

С учетом граничных условий (8-88) и (8-89) выражение (8-85) для тока принимает вид:

$$j = \frac{n_0 \mu_n k_0 T \left[1 - e^{-\frac{eV}{k_0 T}} \right]}{\int_0^{\mathcal{L}} e^{\frac{eU(x)}{k_0 T}} dx}. \quad (8-92)$$

Для вычисления знаменателя дроби произведем замену переменной интегрирования:

$$\int_0^{\mathcal{L}} e^{\frac{eU(x)}{k_0 T}} dx = \int_{U_K - V}^0 e^{\frac{eU(x)}{k_0 T}} \left(\frac{dU}{dx} \right)^{-1} dU. \quad (8-93)$$

Наибольший вклад в величину интеграла будет вносить та область, где значение $U(x)$ максимально, т. е. участок вблизи $x = 0$. С увеличением x функция, стоящая под знаком интеграла, в результате уменьшения $U(x)$ будет резко падать и вкладом ее можно пренебречь. Поэтому для приближенного вычисления интеграла можно вынести за знак интегрирования значение $\left(\frac{dU}{dx} \right)^{-1}$ в точке $x = 0$:

$$\left(\frac{dU}{dx} \right)_{x=0}^{-1} = -\mathcal{E}_s^{-1}. \quad (8-94)$$

На основании (8-94) будем иметь:

$$\int_{U_K - V}^0 e^{\frac{eU(x)}{k_0 T}} \left(\frac{dU}{dx} \right)^{-1} dU = -\mathcal{E}_s^{-1} \frac{k_0 T}{e} \left[1 - e^{\frac{e(U_K - V)}{k_0 T}} \right] \approx \mathcal{E}_s^{-1} \frac{k_0 T}{e} e^{\frac{e(U_K - V)}{k_0 T}}. \quad (8-95)$$

Здесь мы пренебрегли единицей, стоящей в скобках, так как в соответствии с условием (8-90)

$$e \frac{e(U_k - V)}{k_0 T} \gg 1. \quad (8-96)$$

Подставляя в (8-92) выражение (8-95), находим:

$$j = en_0 \mu_n \mathcal{E}_s e^{-\frac{eU_k}{k_0 T}} \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-97)$$

Используя соотношение (8-91), выражение для вольт-амперной характеристики контакта металл — полупроводник можно записать в виде

$$j = en_s \mu_n \mathcal{E}_s \left(e^{\frac{eV}{k_0 T}} - 1 \right) = j_s \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-98)$$

Величина j_s , определяющая значение тока при больших обратных смещениях, в отличие от тока насыщения в диодной теории зависит от приложенного напряжения.

Вольт-амперные характеристики контакта металл — полупроводник для диодной и диффузионной теорий приведены на рис. 8-11 (кривая 1 и 2 соответственно).

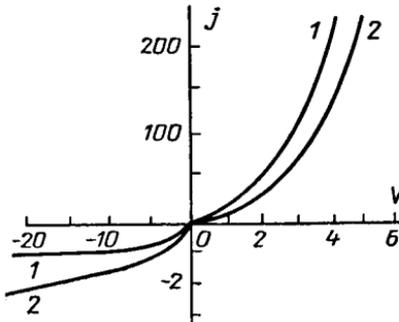


Рис. 8-11. Вольт-амперные характеристики контакта металл — полупроводник.

1 — диодная теория; 2 — диффузионная теория (в относительных единицах)

8-8. КОНТАКТ ЭЛЕКТРОННОГО И ДЫРОЧНОГО ПОЛУПРОВОДНИКОВ

Рассмотрим наиболее простой случай резкого $p-n$ перехода, образованного на границе областей полупроводника с дырочной и электронной проводимостью (рис. 8-12, а).

Пусть акцепторная область полупроводника легирована сильнее, чем донорная ($N_a > N_d$). При возникновении контакта между областями полупроводника с различным типом проводимости начнется взаимная диффузия основных носителей заряда. Электроны будут переходить из n -области в p -область, а диффузия дырок будет идти в противоположном направлении. Возникновение диффузионных потоков приведет к разделению зарядов, вследствие чего появится объемный заряд, положительный в n -области и отрицательный в p -области, и в области контакта полупроводников с разным типом проводимости возникнет электрическое поле, направленное от n -области к p -области, (рис. 8-12, б). Под действием этого поля возникнет дрейф дырок из n -области в p -область и электронов из p - в n -область. Обмен носителями между n - и p -областями будет продолжаться до тех пор, пока дрейфовый ток неосновных носителей не

сравнивается по величине с диффузионным током основных носителей. В результате наступит равновесное состояние, при котором суммарный ток через p - n переход равен нулю. Равновесное состояние будет характеризоваться одинаковым положением уровня Ферми во всей системе, а уровни энергии в приконтактной области, в которой имеется электрическое поле, будут искривлены (рис. 8-12, в). Искривления зон энергии вызовет перераспределение концентрации электронов и дырок.

На рис. 8-12, г изображено распределение основных носителей заряда, дырок p_p и электронов n_n , и неосновных носителей, электронов в p -области n_p и дырок в n -области p_n .

Будем считать, что полупроводник сильно легирован, так что примесные атомы полностью ионизированы. В состоянии термодинамического равновесия

$$p_p n_p = n_n p_n = n_i^2. \quad (8-99)$$

Поскольку p -область сильнее легирована, чем n -область, то $p_p \gg n_n$ и, следовательно, глубина проникновения контактного поля в более низкоомный дырочный полупроводник будет меньше, чем в электронный, т. е. $\mathcal{L}_n > |\mathcal{L}_p|$ (рис. 8-12, д), а вся область объемного заряда будет равна:

$$\mathcal{L}_0 = |\mathcal{L}_p| + \mathcal{L}_n. \quad (8-100)$$

В интервале $-\mathcal{L}_p < x < 0$ отрицательный объемный заряд определяется ионами акцепторной примеси

$$\rho = -eN_a^- = -ep_p, \quad (8-101)$$

а в области $0 < x < \mathcal{L}_n$ — положительными ионами донорной примеси

$$\rho = eN_d^+ = en_n. \quad (8-102)$$

Величина потенциального барьера p - n перехода равна

$$eU_k = \Phi_n - \Phi_p = (\chi_0 + E_c - F_n) - (\chi_0 + E_c - F_p) = F_p - F_n. \quad (8-103)$$

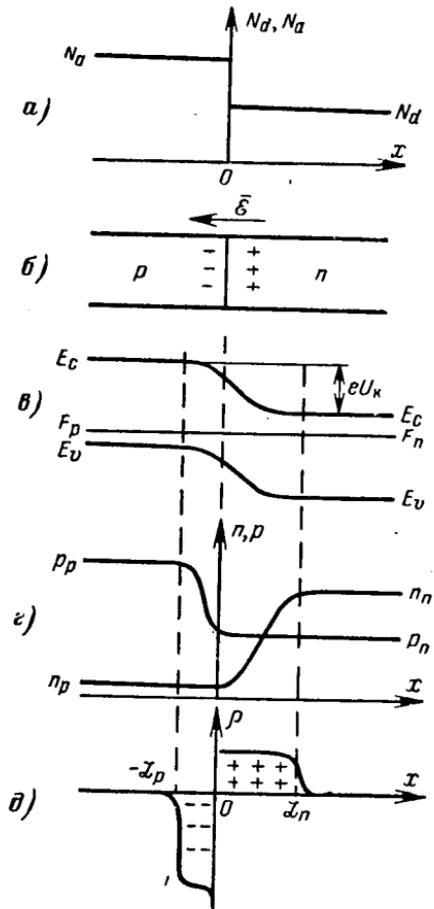


Рис. 8-12. Распределение примесей (а), разделение зарядов и возникновение электрического поля (б), зонная структура (в), распределение концентрации электронов и дырок (г) и распределение объемного заряда (д) в контакте электронного и дырочного полупроводников.

Используя соотношения (8-26) и (8-27), получаем:

$$eU_k = \Delta E_g - k_0 T \ln \frac{N_v}{N_a} - k_0 T \ln \frac{N_c}{N_d} = \Delta E_g - k_0 T \ln \frac{N_c N_v}{N_a N_d}. \quad (8-104)$$

Если учесть, что

$$n_i^2 = N_c N_v e^{-\frac{\Delta E_g}{k_0 T}}, \quad (8-105)$$

то выражение (8-104) примет вид:

$$eU_k = k_0 T \ln \frac{n_n p_p}{n_i^2} \quad (8-106)$$

или

$$eU_k = k_0 T \ln \frac{p_p}{p_n} = k_0 T \ln \frac{n_n}{n_p}. \quad (8-107)$$

Отсюда следует, *чем сильнее легированы обе области полупроводника, тем больше контактная разность потенциалов. Ее максимальное значение*

$$U_{\text{к макс}} = \frac{1}{e} \Delta E_g.$$

На основании формулы (8-107) можно получить следующие соотношения для концентраций основных и неосновных носителей заряда:

$$p_n = p_p e^{-\frac{eU_k}{k_0 T}}; \quad n_p = n_n e^{-\frac{eU_k}{k_0 T}}. \quad (8-108)$$

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 $T = 300^\circ\text{K}$ и значениях концентраций

$$\begin{aligned} N_d = n_n = 10^{16} \text{ см}^{-3}; \quad p_n = 10^{10} \text{ см}^{-3}; \\ N_a = p_p = 10^{14} \text{ см}^{-3}; \quad n_p = 10^{12} \text{ см}^{-3} \end{aligned}$$

Величина потенциального барьера составляет:

$$eU_k = k_0 T \ln p_p/p_n = 0,25 \text{ эв.}$$

Определим глубину проникновения электрического поля в каждую из областей. Уравнение Пуассона для области ($0 < x < L_n$) с учетом соотношения $\rho = en_n$ имеет вид:

$$\frac{d^2 U}{dx^2} = -\frac{en_n}{\epsilon \epsilon_0}. \quad (8-109)$$

Для p -области ($-\mathcal{L}_p < x < 0$), используя равенство $\rho = -e\rho_p$, получаем:

$$\frac{d^2U}{dx^2} = \frac{e\rho_p}{\epsilon\epsilon_0}. \quad (8-110)$$

Граничные условия, отражающие отсутствие поля в глубине материала, запишутся следующим образом:

$$U(\mathcal{L}_n) = 0; \quad \left. \frac{dU}{dx} \right|_{x=\mathcal{L}_n} = 0; \quad (8-111)$$

$$U(-\mathcal{L}_p) = U_k; \quad \left. \frac{dU}{dx} \right|_{x=-\mathcal{L}_p} = 0. \quad (8-112)$$

Согласно соотношению (8-50) этим условиям удовлетворяют следующие выражения для U , являющиеся решением уравнений (8-109) и (8-110):

$$\text{При } 0 < x < \mathcal{L}_n \quad U_n = \frac{en_n}{2\epsilon\epsilon_0} (\mathcal{L}_n - x)^2. \quad (8-113)$$

$$\text{При } -\mathcal{L}_p < x < 0 \quad U_p = U_k - \frac{e\rho_p}{2\epsilon\epsilon_0} (\mathcal{L}_p + x)^2. \quad (8-114)$$

На границе раздела ($x = 0$) потенциалы и значения напряженности поля в обеих областях должны быть одинаковы:

$$U_n(0) = U_p(0); \quad (8-115)$$

$$\left. \frac{dU_n}{dx} \right|_{x=0} = \left. \frac{dU_p}{dx} \right|_{x=0}. \quad (8-116)$$

Применяя условие (8-116) к выражениям (8-113) и (8-114), получаем:

$$n_n \mathcal{L}_n = \rho_p \mathcal{L}_p. \quad (8-117)$$

Это равенство означает, что объемные заряды в обеих областях, прилегающих к p - n переходу, равны между собой. Из уравнения (8-117) нетрудно получить следующие соотношения:

$$\frac{\mathcal{L}_n}{\mathcal{L}_0} = \frac{\rho_p}{n_n + \rho_p}; \quad \frac{\mathcal{L}_p}{\mathcal{L}_0} = \frac{n_n}{n_n + \rho_p}. \quad (8-118)$$

Используя условие (8-115), находим:

$$U_k = \frac{e}{2\epsilon\epsilon_0} (n_n \mathcal{L}_n^2 + \rho_p \mathcal{L}_p^2) = \frac{e}{2\epsilon\epsilon_0} \mathcal{L}_0^2 \frac{n_n \rho_p}{n_n + \rho_p}, \quad (8-119)$$

откуда следует выражение для полной толщины слоя объемного заряда

$$\mathcal{L}_0 = \sqrt{\frac{2\epsilon\epsilon_0}{e} U_k \frac{n_n + \rho_p}{n_n \rho_p}}. \quad (8-120)$$

Из полученной формулы видно, чем выше степень легирования полупроводника, тем меньше величина \mathcal{L}_0 . Если одна из областей легирована значительно сильнее другой (например, $p_p \gg n_n$), то поле практически распространяется только в слабо легированный материал

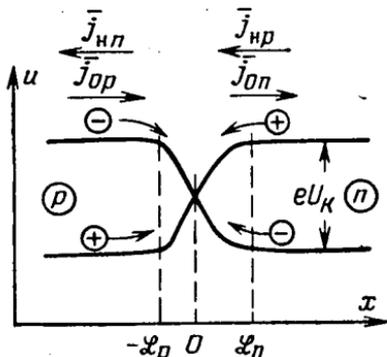


Рис. 8-13. Потенциальные барьеры для дырок и электронов в p - n переходе.

$$\mathcal{L}_0 = \mathcal{L}_n = \sqrt{\frac{2\epsilon\epsilon_0 U_k}{en_n}}. \quad (8-121)$$

Определим падение потенциала в каждой из областей

$$\Delta U_n = U_n(0) = U_k \frac{p_p}{n_n + p_p}; \quad (8-122)$$

$$\begin{aligned} \Delta U_p = U_k - U_p(0) &= U_k - U_n(0) = \\ &= U_k \frac{n_n}{n_n + p_p}. \end{aligned} \quad (8-123)$$

Из формул (8-122) и (8-123) следует, что большая часть падения потенциала приходится на высокоомную область.

На рис. 8-13 схематически изображены потенциальные барьеры для электронов и дырок и показаны направления составляющих тока основных и неосновных носителей из n - и p -областей.

Основные носители при переходе в соседнюю область должны преодолеть потенциальный барьер eU_k . Переход неосновных носителей совершается под действием электрического поля. В состоянии термодинамического равновесия диффузионный поток основных носителей уравновешен дрейфовым током неосновных носителей заряда, и суммарный ток через p - n переход равен нулю.

8-9. ВЫПРЯМЛЕНИЕ НА p - n ПЕРЕХОДЕ

В предыдущем параграфе мы имели дело с p - n переходом, к которому не было приложено внешнее напряжение. Рассмотрим теперь случай, когда на p - n переход подается прямое смещение V (минус к n -области). Высота потенциального барьера снижается на величину eV по сравнению с равновесным состоянием. Соответственно изменяется и толщина запирающего слоя; для ее вычисления нужно вместо формулы (8-120) использовать выражение

$$\mathcal{L} = \sqrt{\frac{2\epsilon\epsilon_0 (U_k - V) n_n + p_p}{e n_n p_p}}. \quad (8-124)$$

Ток неосновных носителей через переход остается практически неизменным, в то время как ток основных носителей заряда возрастает, так как с понижением потенциального барьера большее число электронов может перейти в p -область и большее число дырок — в n -область. В результате этого во внешней цепи будет протекать

ток, равный разности токов основных и неосновных носителей:

$$j = j_0 - j_n. \quad (8-125)$$

В n -области появившиеся неравновесные неосновные носители — дырки Δp создадут в первый момент вблизи контакта положительный объемный заряд. Однако через очень короткое время, определяемое максвелловским временем релаксации (см. § 6-3), этот заряд будет скомпенсирован объемным зарядом основных носителей — электронов, которые под действием электрического поля, созданного неравновесными дырками, будут подтянуты в количестве Δn из глубины n -области, а в n -область электроны поступят из внешней цепи. Таким образом, во всех частях полупроводника будет соблюдаться электрическая нейтральность, но в приконтактной области концентрация электронов и дырок будет повышена на величину $\Delta p = \Delta n$ по сравнению с равновесным состоянием. *Введение в полупроводник неравновесных носителей заряда с помощью p - n перехода при подаче на него прямого смещения называется инжекцией.* Концентрация дырок вблизи контакта будет равна:

$$p_1 = p_n + \Delta p. \quad (8-126)$$

Чтобы вычислить величину p_1 , нужно в равенстве (8-108) вместо U_k использовать значение ($U_k - V$):

$$p_1 = p_n e^{-\frac{e(U_k - V)}{k_0 T}} = p_n e^{\frac{eV}{k_0 T}}. \quad (8-127)$$

Теперь для концентрации неравновесных дырок в n -области получим:

$$\Delta p = p_1 - p_n = p_n \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-128)$$

Аналогичные явления происходят в p -области: сюда из n -области инжектируются неравновесные электроны. Концентрация неравновесных электронов вблизи контакта в p -области увеличивается по сравнению с равновесным значением n_p на величину

$$\Delta n = n_p \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-129)$$

Из найденных выражений следует, что с увеличением прямого смещения на p - n переходе концентрация инжектируемых неосновных носителей резко возрастает, что приводит к сильному росту тока через контакт в прямом направлении.

При обратном подключении батареи (плюс к n -области) потенциальный барьер повышается и увеличивается толщина запирающего слоя. Чтобы найти ее значение, нужно в формулу (8-124) подставить величину обратного напряжения V с отрицательным знаком. Чем сильнее смещен переход, тем меньшее количество основных носителей способно преодолеть возросший потенциальный барьер. Концентрация электронов в p -области по-прежнему определяется равенством (8-129), но значение V здесь также берется с отрицательным знаком.

В соответствии с этим количество основных носителей заряда в приконтактной области уменьшается по сравнению с равновесным состоянием и уменьшается количество неосновных носителей вследствие соблюдения нейтральности. Следовательно, при обратном подключении батареи к p - n переходу имеет место экстракция носителей заряда. Дрейфовый ток через переход практически не изменяется. Поэтому суммарный обратный ток через p - n переход вначале растет с увеличением обратного напряжения, а затем стремится к некоторой величине, называемой током насыщения.

Объемный заряд, сосредоточенный в приконтактной области, изменяется в зависимости от внешнего напряжения, поскольку изменяется толщина запирающего слоя. Благодаря этому электронно-дырочный переход подобно контакту металл — полупроводник обладает некоторой емкостью. Удельную величину ее (на единицу площади) можно приблизительно оценить по формуле

$$C = \frac{\epsilon\epsilon_0}{\mathcal{L}},$$

что с учетом равенства (8-124) приводит к выражению

$$C = \sqrt{\frac{\epsilon\epsilon_0}{2(U_k - V)} \frac{n_n p_p}{n_n + p_p}}. \quad (8-130)$$

Как и в случае контакта металл — полупроводник, емкость p - n перехода при положительном смещении больше, чем при отрицательном.

8-10. ТЕОРИЯ ТОНКОГО p - n ПЕРЕХОДА

Проведем расчет вольт-амперной характеристики p - n перехода. Для упрощения задачи примем ряд допущений.

1. Будем рассматривать тонкий переход, когда носители пролетают через слой объемного заряда без рекомбинации. Это означает, что толщина запирающего слоя значительно меньше диффузионной длины:

$$\mathcal{L}_0 \ll L_D. \quad (8-131)$$

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 подается внешнее смещение, сделаны антизапорными и находятся далеко от перехода, так что неравновесные носители рекомбинируют полностью, не доходя до них. Падением напряжения на этих контактах также можно пренебречь и считать, что все напряжение батареи приложено к p - n переходу.

4. Ловушки и поверхностная рекомбинация отсутствуют, и уменьшение концентрации носителей заряда происходит только в результате рекомбинации в объеме полупроводника.

Чтобы рассчитать вольт-амперную характеристику p - n перехода, нужно найти закон изменения концентрации свободных носителей, решив уравнения непрерывности для дырок и электронов (см. § 7-1), которые имеют вид:

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial j_p}{\partial x} - \frac{\Delta p}{\tau_p}; \quad (8-132)$$

$$\frac{\partial n}{\partial t} = \frac{1}{e} \frac{\partial j_n}{\partial x} - \frac{\Delta n}{\tau_n}. \quad (8-133)$$

Рассмотрим дырочный ток в p - и n -области, считая, что значения коэффициента диффузии и подвижности дырок в обеих областях одинаковы (что, строго говоря, не совсем верно). Полный ток дырок определяется дрейфовой и диффузионной составляющими

$$j = e p \mu_p \mathcal{E} - e D_p \frac{dp}{dx}, \quad (8-134)$$

где \mathcal{E} — напряженность электрического поля. Концентрация дырок в p -области равна $p = p_p + \Delta p$. Здесь $\Delta p = \Delta n$ — количество дырок, подтянувшихся к контакту для нейтрализации заряда электронов, инжектированных из электронного полупроводника. Мы рассматриваем сильно легированный полупроводник, когда величина p_p очень велика ($p_p \gg \Delta p$), и даже при малых \mathcal{E} дрейфовая составляющая тока значительно превосходит диффузионную, поэтому для p -области можно написать:

$$j_p^{(p)} = e p \mu_p \mathcal{E}. \quad (8-135)$$

В противоположность этому в n -области имеем $p = p_n + \Delta p'$, где $\Delta p'$ — количество дырок, инжектированных из p -области. Величина p_n мала, и диффузионная составляющая тока превышает дрейфовую, поэтому

$$j_p^{(n)} = -e D_p \frac{dp}{dx}. \quad (8-136)$$

Так как мы приняли, что в запирающем слое не происходит рекомбинации, дырочные токи в обеих областях на границе запирающего слоя одинаковы:

$$j_p^{(p)}|_{x=-x_p} = j_p^{(n)}|_{x=x_n}. \quad (8-137)$$

Уравнение непрерывности для дырок в n -области для стационарного случая ($\partial p/\partial t = 0$) с учетом (8-136) запишется следующим образом:

$$D_p \frac{d^2 p}{dx^2} - \frac{p - p_n}{\tau_p} = 0. \quad (8-138)$$

Используя соотношение $L_p^2 = D_p \tau_p$, получаем:

$$\frac{d^2 p}{dx^2} - \frac{\Delta p'}{L_p^2} = 0. \quad (8-139)$$

Общее решение такого уравнения имеет вид:

$$\Delta p' = A e^{-\frac{x}{L_p}} + B e^{\frac{x}{L_p}}. \quad (8-140)$$

Так как при движении в глубь полупроводника ($x \rightarrow \infty$) концентрация неравновесных дырок спадает до нуля, константа B должна быть принята равной нулю. Тогда

$$p = p_n + \Delta p' = p_n + A e^{-\frac{x}{L_p}}. \quad (8-141)$$

Для определения константы A воспользуемся выражением для концентрации носителей на границе запертого слоя, аналогичным равенству (8-127):

$$p(\mathcal{L}_n) = p_n e^{\frac{eV}{k_0 T}}.$$

Учитывая это условие, из (8-141) при $x = \mathcal{L}_n$ находим:

$$A = p_n \left(e^{\frac{eV}{k_0 T}} - 1 \right) e^{\frac{\mathcal{L}_n}{L_p}}. \quad (8-142)$$

Окончательно закон изменения концентрации дырок в n -области ($x \geq \mathcal{L}_n$, рис. 8-12, ε) принимает вид:

$$p(x) = p_n + p_n \left(e^{\frac{eV}{k_0 T}} - 1 \right) e^{-\frac{x - \mathcal{L}_n}{L_p}}. \quad (8-143)$$

Для дырочного тока согласно (8-136) получим:

$$j_p^{(n)} = \frac{e D_p p_n}{L_p} \left(e^{\frac{eV}{k_0 T}} - 1 \right) e^{-\frac{x - \mathcal{L}_n}{L_p}}. \quad (8-144)$$

Значение этой составляющей при $x = \mathcal{L}_n$ равно:

$$j_p^{(n)} |_{x = \mathcal{L}_n} = \frac{e D_p p_n}{L_p} \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-145)$$

С помощью аналогичного расчета нетрудно получить выражение для электронной составляющей тока в p -области при $x = -\mathcal{L}_p$, подобное соотношению (8-145):

$$j_p^{(p)} |_{x = -\mathcal{L}_p} = \frac{e D_n n_p}{L_n} \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-146)$$

Принимая во внимание условие (8-137), для общего тока через контакт можно использовать формулу

$$j = j_p^{(n)} |_{x = \mathcal{L}_n} + j_p^{(p)} |_{x = -\mathcal{L}_p}. \quad (8-147)$$

Отсюда на основании выражений (8-145) и (8-146) получаем выражение для вольт-амперной характеристики p - n перехода

$$j = e \left(\frac{D_n n_p}{L_n} + \frac{D_p p_n}{L_p} \right) \left(e^{\frac{eV}{k_0 T}} - 1 \right) = j_s \left(e^{\frac{eV}{k_0 T}} - 1 \right), \quad (8-148)$$

где ток насыщения j_s равен:

$$j_s = j_{ns} + j_{ps} = e \left(\frac{D_n n_p}{L_n} + \frac{D_p p_n}{L_p} \right) = en_i \left(\frac{D_n}{L_n p_p} + \frac{D_p}{L_p n_n} \right) = e \left(\frac{n_p L_n}{\tau_n} + \frac{p_n L_p}{\tau_p} \right) \quad (8-149)$$

Вид вольт-амперной характеристики показан на рис. 8-14.

При прямом смещении ток экспоненциально возрастает, а с ростом отрицательного напряжения стремится к j_s . Выпрямляющие свойства p - n перехода тем лучше, чем меньше ток насыщения. Его величина уменьшается с ростом концентрации основных носителей n_n и p_p (т. е. с увеличением степени легирования) и времени жизни неосновных носителей. Повышение температуры приводит к росту собственной концентрации n_i и значения j_s .

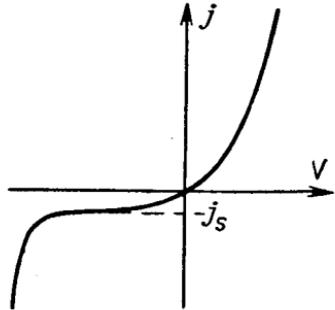


Рис. 8-14. Вольт-амперная характеристика p - n перехода.

При дальнейшем увеличении обратного напряжения в некоторый момент происходит пробой p - n перехода. Энергии электрического поля оказывается достаточно для разрыва валентных связей, и ток через переход резко возрастает (з и н е р о в с к и й п р о б о й).

Другой механизм пробоя связан с тем, что свободные носители могут создаваться за счет кинетической энергии частицы, полученной ею от электрического поля при прохождении запирающего слоя. Возникающие свободные носители в свою очередь ускоряются полем и также могут принять участие в генерации электронно-дырочных пар. Происходит лавинное нарастание обратного тока, благодаря чему этот тип пробоя назван л а в и н н ы м.

Рассмотрим теперь соотношение токов, текущих через p - n переход. В соответствии с формулами (8-144) и (8-146) отношение электронной и дырочной составляющих токов через контакт равно:

$$\frac{j_n}{j_p} = \frac{j_{ns}}{j_{ps}} = \frac{D_n n_p L_p}{D_p p_n L_n} = \frac{\mu_n n_n L_p}{\mu_p p_p L_n} = \frac{\sigma_n L_p}{\sigma_p L_n} \quad (8-150)$$

Из полученного равенства следует, что отношение электронного тока к дырочному определяется главным образом отношением концентраций основных носителей в n - и p -областях. Если степень легирования обеих частей полупроводника приблизительно одинакова ($\sigma_n = \sigma_p$), то такой p - n переход инжектирует в равной степени электроны и дырки и общий ток определяется выражением

$$j = (j_{ns} + j_{ps}) \left(e^{\frac{eV}{k_0 T}} - 1 \right). \quad (8-151)$$

В том случае, когда p -область легирована значительно сильнее n -области ($\sigma_p \gg \sigma_n$), получаем: $j_{ps} \gg j_{ns}$, и через переход течет в основном дырочный ток. При $\sigma_n \gg \sigma_p$, напротив, основной вклад в общий ток вносит электронная составляющая.

Перейдем к рассмотрению распределения концентраций носителей заряда и токов в p - n переходе. Концентрация дырок в n -области ($x \geq \mathcal{L}_n$) определяется равенством (8-143). Для электронов в p -области ($x \leq -\mathcal{L}_p$) нетрудно получить аналогичное выражение

$$n(x) = n_p + n_p \left(e^{\frac{eV}{k_0 T}} - 1 \right) e^{\frac{x + \mathcal{L}_p}{L_n}}. \quad (8-152)$$

Концентрации основных носителей в обеих областях можно считать неизменными в соответствии с условиями

$$p_p \gg \Delta p; \quad n_n \gg \Delta n.$$

Дырочная составляющая тока в n -области выражается равенством (8-144). По аналогии с этим для электронного тока в p -области имеем:

$$j_n^{(p)} = j_{ns} \left(e^{\frac{eV}{k_0 T}} - 1 \right) e^{\frac{x + \mathcal{L}_p}{L_n}}. \quad (8-153)$$

Сумма электронного и дырочного тока при любом значении x постоянна

$$j = j_p^{(n)}(x) + j_n^{(n)}(x) = j_n^{(p)}(x) + j_p^{(p)}(x) = \text{const}. \quad (8-154)$$

Кривые распределения концентраций носителей и токов через p - n переход, построенные в соответствии с приведенными формулами, показаны на рис. 8-15.

При прямом смещении перехода концентрация неосновных носителей у контакта повышается (явление инжекции). В случае подачи обратного напряжения имеет место вытягивание неосновных носителей из приконтактной области (явление экстракции).

Так как на расстоянии диффузионной длины от контакта концентрация неосновных носителей заряда отличается от равновесной, эта область характеризуется квазиуровнями Ферми. Чтобы определить изменение положения квазиуровней Ферми, рассмотрим концентрацию электронов в точке $x = -\mathcal{L}_p$. Согласно выражению (8-152)

$$n(-\mathcal{L}_p) = n_p e^{\frac{eV}{k_0 T}} = N_c e^{\frac{F_n(-\mathcal{L}_p) - E_c}{k_0 T}}, \quad (8-155)$$

где $F_n(-\mathcal{L}_p)$ — квазиуровень Ферми для электронов при $x = -\mathcal{L}_p$.

Равновесная концентрация электронов в p -области равна:

$$n_p = N_c e^{\frac{F_p - E_c}{k_0 T}}. \quad (8-156)$$

Подставляя равенство (8-156) в выражение (8-62), находим:

$$F_p = F_n(-\mathcal{L}_p) - eV, \quad (8-157)$$

т. е. на расстоянии, которое определяется диффузионной длиной L_n , квазиуровень Ферми для электронов в p -области смещается на

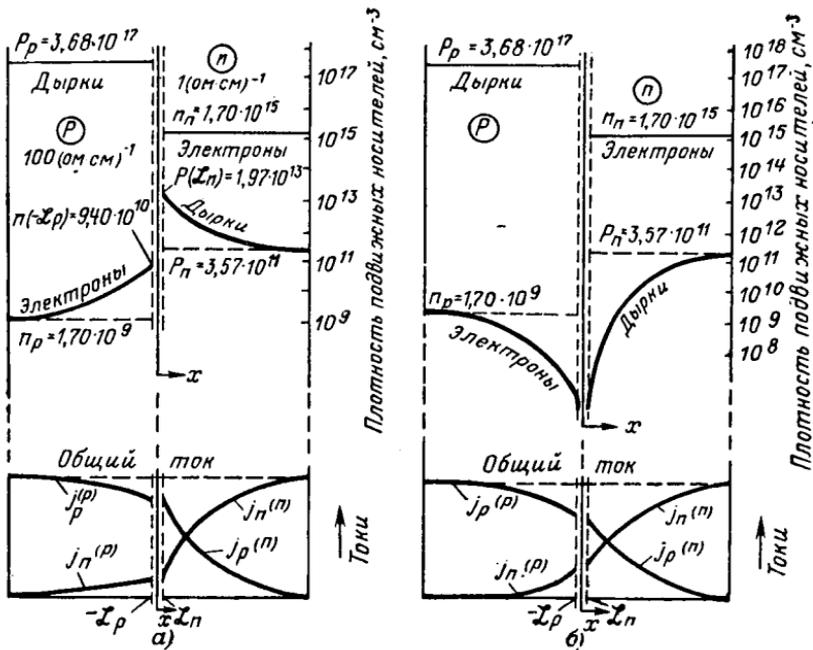


Рис. 8-15. Концентрации носителей заряда и токов в p - n переходе при прямом (а) и обратном (б) смещении.

величину eV . Аналогично этому происходит изменение положения квазиуровня Ферми для дырок в n -области (рис. 8-16).

При прямом смещении p - n перехода концентрация неосновных носителей у контакта возрастает. В соответствии с этим дырочный

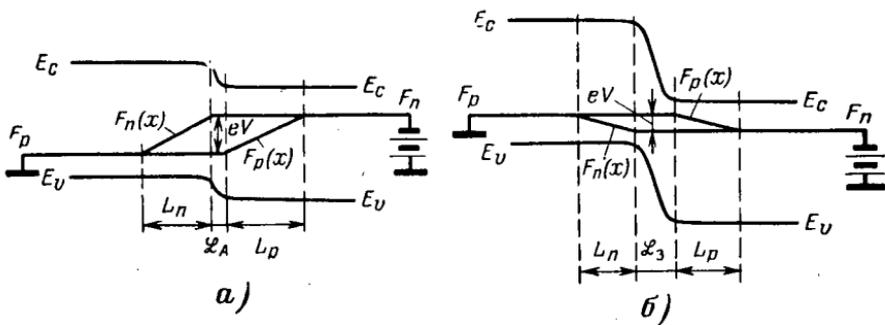


Рис. 8-16. Квазиуровень Ферми в p - n переходе при прямом (а) и обратном (б) смещении.

квазиуровень Ферми в n -области приближается к верхнему краю валентной зоны, а квазиуровень Ферми для электронов в p -области — к дну зоны проводимости (рис. 8-16, а). В случае подачи отрицатель-

ного напряжения происходит уменьшение количества неосновных носителей вблизи контакта и квазиуровни Ферми в приконтактной области удаляются от соответствующих зон (рис. 8-16, б).

Глава девятая

ПОВЕРХНОСТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

9-1. ПОВЕРХНОСТНЫЙ ПОТЕНЦИАЛ

Для бесконечного кристалла в случае сильно связанных электронов, как следует из решения уравнения Шредингера (гл. 2), достаточно рассмотреть две зоны разрешенных значений энергии — зону проводимости и валентную зону. В ограниченном кристалле, кроме зон разрешенных энергий, на его поверхности вследствие разрыва периодичности кристаллической решетки возникают поверхностные состояния, энергетические уровни которых расположены в запрещенной зоне. Эти уровни теоретически были предсказаны Таммом и носят название у р о в н е й Т а м м а. Концентрация таммовских уровней определяется концентрацией поверхностных атомов и составляет величину 10^{15} см^{-2} .

На реальной поверхности полупроводника, находящегося в контакте с окружающей атмосферой, поверхностные состояния возникают также и вследствие адсорбции примесей (атомов или ионов). Такие примеси, как, например, кислород и окись углерода, выступают в качестве доноров электронов или ведут себя как акцепторы электронов (табл. 9-1).

Таблица 9-1

Изменение знака заряда на поверхности полупроводника при адсорбции примесей

Полупроводник	O ₂	CO	CO ₂	H ₂ O	Полупроводник	O ₂	CO	CO ₂	H ₂ O
Cu ₂ O	-	+	-	+	NiO	-	+	-	-
CuO	-	+	-	+	Ge	-	+	+	+

В зависимости от сил, связывающих адсорбирование частицы с поверхностью полупроводника, имеет место физическая или химическая адсорбция (хемосорбция). Физическая адсорбция определяется силами электростатического происхождения (силы Ван-дер-Ваальса, силы электрического изображения), и энергия связи в этом случае составляет величину порядка 0,01—0,1 эв. Химическая адсорбция возникает, когда адсорбированные молекулы связаны с полупроводником силами обменного типа. В силу этого энергия связи при хемосорбции значительна и достигает 1 эв.

На поверхности полупроводника могут также находиться центры рекомбинации и центры прилипания. Энергетические уровни всех этих поверхностных состояний обычно располагаются в запрещенной зоне значительно ниже дна зоны проводимости или выше валентной зоны.

Поверхностный заряд притягивает носители из объема полупроводника в приповерхностную область. В результате этого возникает двойной заряженный слой. В металле, где концентрация свободных электронов в единице объема 10^{22} см^{-3} , нейтрализация поверхностного заряда происходит на расстоянии 10^{-7} см . Такой тонкий слой не может существенно сказываться на свойствах всего кристалла. У полупроводников концентрация свободных носителей значительно меньше, поэтому область пространственного заряда проникает достаточно глубоко в кристалл. Например, у германия, у которого концентрация поверхностных уровней в единице объема составляет величину 10^{11} см^{-2} , при концентрации носителей в 10^{15} см^{-3} нейтрализация поверхностного заряда происходит на расстоянии 10^{-4} см , а в собственном германии уже на расстоянии 0,1 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_s , расположенные ниже уровня Ферми в объеме материала (рис. 9-1). Электроны зоны проводимости будут стремиться заполнить поверхностные состояния. Если плотность поверхностных состояний N_s , то при заполнении всех поверхностных состояний на каждой единице площади поверхности возникнет отрицательный заряд, равный $-eN_s$. В приповерхностном слое при этом появится обедненный электронами слой, обладающий положительным пространственным зарядом, и возникнет электрическое поле, направленное к поверхности полупроводника. Наличие электрического поля в приповерхностном слое полупроводника приводит к изгибу энергетических зон вблизи поверхности полупроводника. Если через $e\varphi_s$ обозначить величину изгиба зоны проводимости на поверхности, то φ_s — поверхностный потенциал.

В приповерхностном слое электронного полупроводника в зависимости от положения границы его зон относительно уровня Ферми в общей сложности можно выделить две области. Область I, так называемая область обеднения, для которой проводи-

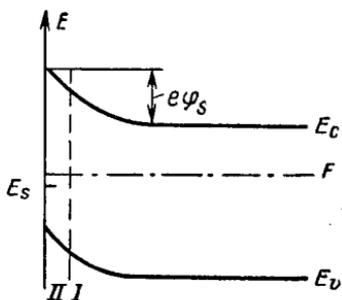


Рис. 9-1. Энергетическая диаграмма приповерхностной области полупроводника *n*-типа. I — область обеднения; II — область инверсии.

мость меньше, чем в объеме полупроводника, и область II, так называемый и н в е р с и о н н ы й с л о й, для которого имеет место изменение типа проводимости (рис. 9-1).

Образование инверсионного слоя зависит от степени легирования образца. Для образования инверсионного слоя у полупроводника, проводимость которого сильно отклоняется от собственной, требуется большое значение поверхностного потенциала φ_s .

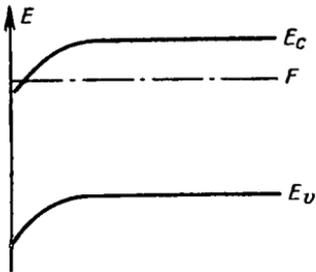


Рис. 9-2. Энергетическая диаграмма приповерхностной области полупроводника *n*-типа в случае обогащения.

Если в результате наличия поверхностных состояний зоны электронного полупроводника изгибаются вниз, то образуется область обогащения (рис. 9-2).

В этом случае концентрация основных носителей у поверхности больше, чем в объеме.

У полупроводников *p*-типа обеднение наблюдается в том случае, если на поверхности захватываются дырки и зоны изгибаются вниз, а при захвате электронов имеет место обогащение, при котором зоны изгибаются вверх.

Рассмотрим теорию приповерхностного объемного заряда для электронного полупроводника. Будем считать, что концентрация донорной примеси равна N_d и имеет место полная ее ионизация. Для определения электростатического потенциала в приповерхностной области этого полупроводника необходимо решить уравнение Пуассона

$$\frac{d^2\varphi}{dx^2} = -\frac{1}{\epsilon\epsilon_0} \rho. \quad (9-1)$$

В объеме полупроводника суммарный объемный заряд равен нулю и, следовательно,

$$n_0 = N_d + p_0. \quad (9-2)$$

Вблизи поверхности, где есть электрическое поле, значения концентрации электронов и дырок будут другими, чем в глубине полупроводника (рис. 9-3), поэтому $n - p - N_d \neq 0$ и объемный заряд будет равен:

$$\rho = -e[n(x) - n_0] + e[p(x) - p_0]. \quad (9-3)$$

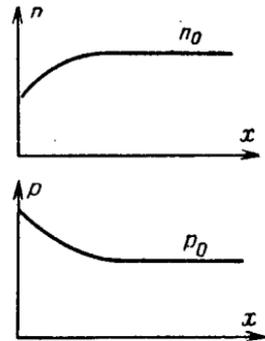


Рис. 9-3. Распределение концентрации электронов и дырок в глубине и в приповерхностном слое полупроводника *n*-типа.

Для невырожденного полупроводника концентрации электронов и дырок в приповерхностной области определяются распределением Больцмана. Их можно выразить через поверхностный потенциал следующим образом:

$$n = n_0 e^{-\frac{e\varphi}{k_0 T}}; \quad p = p_0 e^{\frac{e\varphi}{k_0 T}}. \quad (9-4)$$

Теперь объемный заряд вблизи поверхности с учетом выражения (9-4) можно записать в виде

$$\rho = -en_0 \left(e^{-\frac{e\varphi}{k_0 T}} - 1 \right) + e\rho_0 \left(e^{\frac{e\varphi}{k_0 T}} - 1 \right). \quad (9-5)$$

Для удобства решения этого уравнения введем следующие переменные:

$$\Psi = \frac{e\varphi}{k_0 T}; \quad \gamma = \frac{n_0}{n_i} = \frac{n_i}{p_0}; \quad L^2 = \frac{\varepsilon \varepsilon_0 k_0 T}{2e^2 n_i}. \quad (9-6)$$

Подставив выражение для плотности объемного заряда в уравнение Пуассона, получим:

$$2 \frac{d^2 \Psi}{dx^2} = L^{-2} [\gamma (e^{-\Psi} - 1) - \gamma^{-1} (e^{\Psi} - 1)]. \quad (9-7)$$

Чтобы решить это уравнение, умножим его правую и левую части на $d\Psi/dx$ и используем тождество

$$2 \frac{d\Psi}{dx} \frac{d^2 \Psi}{dx^2} = \frac{d}{dx} \left(\frac{d\Psi}{dx} \right)^2. \quad (9-8)$$

Тогда, умножая левую и правую части на dx и интегрируя по Ψ , находим:

$$\int d \left(\frac{d\Psi}{dx} \right)^2 + C' = L^{-2} \int_0^{\Psi} [\gamma (e^{-\Psi} - 1) - \gamma^{-1} (e^{\Psi} - 1)] d\Psi, \quad (9-9)$$

откуда будем иметь:

$$\left(\frac{d\Psi}{dx} \right)^2 = L^{-2} F^2(\Psi, \gamma) + C, \quad (9-10)$$

где

$$F^2(\Psi, \gamma) = \gamma(1 - e^{-\Psi}) + \gamma^{-1}(1 - e^{\Psi}) + \Psi(\gamma^{-1} - \gamma). \quad (9-11)$$

Константу C можно определить из граничного условия: в объеме полупроводника

$$\text{при } x \rightarrow \infty \quad \Psi = 0 \text{ и } \frac{d\Psi}{dx} = 0. \quad (9-12)$$

При этом условии $F(\Psi) = 0$, а следовательно, и $C = 0$. Тогда из (9-10) найдем:

$$\frac{d\Psi}{dx} = \pm L^{-1} F(\Psi, \gamma). \quad (9-13)$$

Так как в рассматриваемом случае зоны изогнуты вверх (рис. 9-1), то $e\varphi > 0$ и растёт, а следовательно, $\frac{d\Psi}{dx} > 0$. С учетом этого уравнение (9-13) можно записать в виде

$$\frac{d\Psi}{dx} = L^{-1} F(\Psi, \gamma). \quad (9-14)$$

Для определения величины изгиба зон у свободной поверхности

$$\varphi_s = \frac{k_0 T}{e} \Psi_s \quad (9-15)$$

можно воспользоваться условием электронейтральности, согласно которому положительный заряд в приповерхностном слое Q_0 равен отрицательному заряду на поверхностных уровнях Q_s , т. е.

$$Q_s = Q_0. \quad (9-16)$$

Поскольку число электронов на поверхностных уровнях

$$n_s = \frac{N_s}{e \frac{E_s - F + e\varphi_s}{k_0 T} + 1}, \quad (9-17)$$

то полный отрицательный заряд на поверхностных уровнях равен:

$$Q_s = en_s = \frac{eN_s}{e \frac{E_s - F + e\varphi_s}{k_0 T} + 1}. \quad (9-18)$$

Полный положительный заряд в приповерхностном слое можно определить, если учесть выражения (9-1), (9-6), (9-8) и (9-12):

$$\begin{aligned} Q_0 &= \int_0^{\infty} \rho(x) dx = -\frac{\varepsilon\varepsilon_0 k_0 T}{e} \int_0^{\infty} \frac{d^2\Psi}{dx^2} dx = -\frac{\varepsilon\varepsilon_0 k_0 T}{e} \left. \frac{d\Psi}{dx} \right|_0^{\infty} = \\ &= \frac{\varepsilon\varepsilon_0 k_0 T}{e} \left. \frac{d\Psi}{dx} \right|_{x=0}. \end{aligned} \quad (9-19)$$

Так как $\frac{\varepsilon\varepsilon_0 k_0 T}{e} = 2en_i L^2$, то на основании уравнения (9-14) заряд Q_0 можно записать в виде

$$Q_0 = 2en_i LF(\Psi_s, \gamma). \quad (9-20)$$

На основании выражений (9-16), (9-18) и (9-20) можно получить равенство следующего вида:

$$2n_i LF(\Psi_s, \gamma) = n_s, \quad (9-21)$$

из которого можно найти φ_s .

Для нахождения распределения потенциала в области объемного заряда необходимо согласно уравнению (9-14) вычислить интеграл

$$x(\Psi) = \int_0^{\Psi} \frac{d\Psi}{L^{-1}F(\Psi, \gamma)}, \quad (9-22)$$

который в общем случае не решается в элементарных функциях. Для нахождения $\Psi(x)$ обычно решают задачу для трех областей пространственного заряда — для области обеднения, инверсионного слоя и области обогащения.

9-2. ЭФФЕКТ ПОЛЯ

Избыточный заряд на поверхности полупроводника можно создать, используя образец как одну из обкладок конденсатора (рис.9-4). Если к такому конденсатору приложить разность потенциалов,

то на поверхности полупроводника и металла возникнут разноименные заряды, величину которых $Q = CV$ можно определить, измеряя емкость конденсатора на единицу поверхности C и приложенное напряжение V . Поверхностные заряды, кроме того, создаются и за счет контактной разности потенциалов, но если приложенное напряжение значительно больше контактной разности потенциалов, то ее можно не учитывать. Поверхностный заряд, возникший в металле, локализуется непосредственно на его поверхности, а заряд в полупроводнике распределяется между поверхностными уровнями и приповерхностным слоем. В результате этого происходит изменение проводимости полупроводника за счет носителей заряда, возникших в приповерхностном слое. Изменение проводимости в приповерхностной области полупроводника под действием электрического поля называют **э ф ф е к т о м п о л я**.

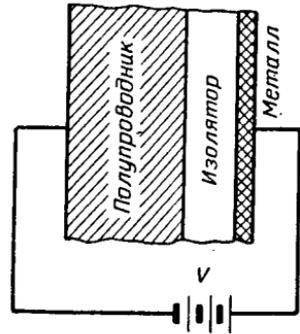


Рис. 9-4. Принцип измерения поверхностной проводимости при эффекте поля.

Избыточная концентрация электронов в приповерхностном слое, отнесенная к единице поверхности полупроводника, с учетом уравнений (9-3), (9-4) и (9-14) может быть определена в виде

$$\begin{aligned} \Delta N &= \int_0^{\infty} [n(x) - n_0] dx = n_0 \int_0^{\infty} (e^{-\Psi} - 1) dx = \\ &= n_0 \int_0^{\Psi_s} (e^{-\Psi} - 1) \left(\frac{d\Psi}{dx}\right)^{-1} d\Psi = n_0 L F(\Psi_s, \gamma), \end{aligned} \quad (9-23)$$

где

$$F(\Psi_s, \gamma) = \int_0^{\Psi_s} \frac{e^{-\Psi} - 1}{F(\Psi, \gamma)} d\Psi. \quad (9-24)$$

Аналогично избыточная концентрация дырок равна:

$$\begin{aligned} \Delta P &= \int_0^{\infty} [p(x) - p_0] dx = p_0 \int_0^{\infty} (e^{\Psi} - 1) dx = p_0 \int_0^{\Psi_s} (e^{\Psi} - 1) \left(\frac{d\Psi}{dx}\right)^{-1} d\Psi = \\ &= p_0 L \Phi(\Psi_s, \gamma), \end{aligned} \quad (9-25)$$

где

$$\Phi(\Psi_s, \gamma) = \int_0^{\Psi_s} \frac{e^{\Psi} - 1}{F(\Psi, \gamma)} d\Psi. \quad (9-26)$$

Используя функции $F(\Psi_s, \gamma)$ и $\Phi(\Psi_s, \gamma)$, для которых составлены таблицы, и зная подвижности электронов μ_n и дырок μ_p , можно рас-

считать избыточную проводимость $\Delta\sigma_s$ в зависимости от величины поверхностного потенциала по формуле

$$\Delta\sigma_s = e\mu_n \Delta N + e\mu_p \Delta P. \quad (9-27)$$

Здесь предполагается, что подвижность носителей в приповерхностном слое такая же, как в объеме. В действительности же в приповерхностном слое она меньше вследствие добавочного рассеяния носителей на поверхности. Рассеяние на поверхности становится существенным, когда избыточные носители заряда в основном сосредото-

члены в слое толщиной порядка длины свободного пробега носителей в объеме. В этом случае рассеяние на поверхности может заметно снизить эффективную длину свободного пробега и уменьшить подвижность.

Ход теоретической зависимости поверхностной проводимости и плотности приповерхностного заряда Q_0 показан на рис. 9-5. При отрицательных значениях $e\varphi_s$ в приповерхностном слое скапливается большое число электронов, что сильно увеличивает значение поверхностной проводимости. При малых отрицательных значениях $e\varphi_s$ происходит убывание концентрации электронов и увеличение концентрации дырок, причем концентрация электронов убывает быстрее, чем рост концентрации дырок, поэтому поверхностная проводимость уменьшается. При $e\varphi_s = 0$ искривление энергетических зон отсутствует и соответственно отсутствует поверх-

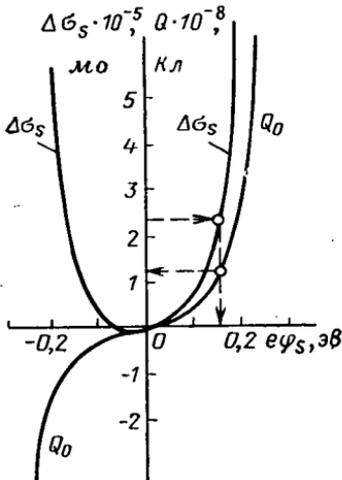


Рис. 9-5. Теоретические кривые поверхностной проводимости $\Delta\sigma_s$ и плотности приповерхностного заряда Q_0 в зависимости от высоты барьера $e\varphi_s$ для германия при комнатной температуре ($N_a = 3,4 \cdot 10^{13} \text{ см}^{-3}$).

стная проводимость. В точке, соответствующей минимуму кривой, снижение поверхностной проводимости за счет уменьшения количества электронов компенсируется увеличением проводимости за счет роста концентрации дырок. При дальнейшем увеличении поверхностного потенциала поверхностная проводимость возрастает за счет повышения концентрации дырок. В результате этого изменяется тип поверхностной проводимости с электронного на дырочный, т. е. имеет место инверсия типа поверхностной проводимости.

Измеряя значение $\Delta\sigma_s$ и используя зависимость Q_0 , можно определить величину поверхностного потенциала. На рис. 9-5 стрелками показано, как, зная $\Delta\sigma_s$, можно определить Q_0 и $e\varphi_s$. Ситуация, однако, осложняется тем обстоятельством, что обычно неизвестно значение $\Delta\sigma_s$ на свободной поверхности, т. е. в отсутствие электрического поля. Поэтому практически поступают следующим образом: сначала строят экспериментальную зависимость $\Delta\sigma_s$ от полного

заряда $Q = CV$; затем определяют поверхностную проводимость в точке минимума; дальше отсчитывают изменение поверхностной проводимости $\Delta\sigma_s$ от $\Delta\sigma_{\min}$. Зная разность $\Delta\sigma_s - \Delta\sigma_{\min}$, можно по теоретической кривой Q_0 определить величину $e\varphi_s$. После этого, сравнивая теоретическую зависимость приповерхностного заряда $Q_0(\varphi_s)$ и экспериментальную зависимость полного заряда $Q(\varphi_s)$, находят захваченный на поверхностных уровнях заряд по формуле:

$$Q_s(\varphi_s) = Q(\varphi_s) - Q_0(\varphi_s). \quad (9-28)$$

Знание зависимости захваченного заряда от φ_s позволяет в принципе рассчитать энергетическое положение и концентрацию поверхностных уровней.

Таким образом, из измерений эффекта поля в стационарном случае возможно определить тип поверхностной проводимости, изгиб зон на поверхности кристалла в зависимости от приложенного напряжения, найти величину заряда в приповерхностной области и заряда, сосредоточенного на поверхности, выявить энергетическое положение поверхностных уровней и определить их концентрацию.

Из измерений эффекта поля можно также определить подвижность носителей в приповерхностном слое, которую в дальнейшем для краткости будем называть подвижностью при эффекте поля $\mu_{э.п}$. Она равна отношению удельной поверхностной проводимости $\Delta\sigma_s$ к полному индуцированному заряду $Q = CV_m$:

$$\mu_{э.п} = \frac{\Delta\sigma_s}{CV_m}, \quad (9-29)$$

где C — удельная емкость системы металл — полупроводник; V_m — напряжение между обкладками конденсатора.

При больших значениях поверхностного потенциала, когда, поверхностная проводимость определяется носителями одного знака, например электронами, $\Delta\sigma_s = e\mu_n\Delta N$, подвижность при эффекте поля равна:

$$\mu_{э.п} = \mu_n \frac{e\Delta N}{Q}. \quad (9-30)$$

Вычислив $\mu_{э.п}$, можно затем определить число носителей заряда, которые вносят вклад в поверхностную проводимость. Очевидно, что в отсутствие процессов захвата индуцированных носителей заряда подвижность при эффекте поля равна холловской подвижности носителей.

$$\mu_{э.п} = \mu_n. \quad (9-31)$$

Как показывают эксперименты, у большинства полупроводников на поверхности имеются два типа уровней с различными временами захвата носителей. Они возникают в результате наличия на поверхности полупроводника, в частности Ge и Si, окисного слоя. Кривая релаксации поверхностной проводимости такого образца представлена на рис. 9-6.

В первый момент времени при включении электрического поля, когда индуцированные носители еще не успели захватиться поверхностными уровнями, поверхностная проводимость имеет максимальное значение. Затем по истечении малого времени (10^{-8} — 10^{-4} сек) происходит уменьшение поверхностной проводимости до величины $\Delta\sigma_1$ и далее медленное снижение ее до величины $\Delta\sigma_2$. Из кривой релаксации можно определить величину заряда на уровнях с большим и малым временами захвата, которые будем называть медленные и быстрые состояния, как отношения $\frac{\Delta\sigma_1 - \Delta\sigma_2}{\Delta\sigma_{\text{макс}}}$ и $\frac{\Delta\sigma_{\text{макс}} - \Delta\sigma_1}{\Delta\sigma_{\text{макс}}}$ соответственно.

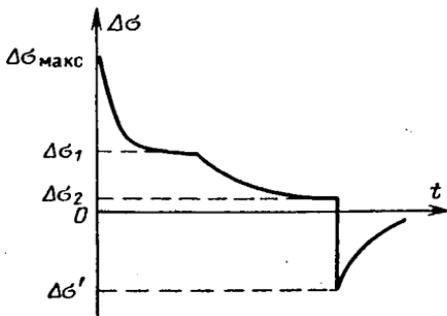


Рис. 9-6. Кривая релаксации поверхностной проводимости.

которых достигает величины 10^{14} — 10^{15} см^{-2} , расположенные на поверхности слоя окиси (4), находящейся на полупроводнике. Большие времена захвата для медленных состояний, изменяющиеся от 10^{-3} сек до нескольких часов, обусловлены тем, что при переходе электрона из полупроводника на эти уровни ему необходимо пройти через сравнительно толстый слой окиси, являющийся диэлектриком. Уровни быстрых состояний находятся непосредственно на поверхности полупроводника под слоем окиси. Плотность их меняется от 10^{11} до 10^{12} см^{-2} в зависимости от обработки поверхности.

Если в некоторый момент времени произвести выключение постоянного электрического поля, создаваемого разностью потенциалов V_m , то между полупроводниковой и металлической пластинами останется напряжение, равное разности потенциалов, которая соответствует захваченному на поверхностные уровни объемному заряду. Полярность этого напряжения V_m' обратна полярности внешнего напряжения V_m , и отношение V_m'/V_m характеризует отношение захваченного и полного объемного зарядов.

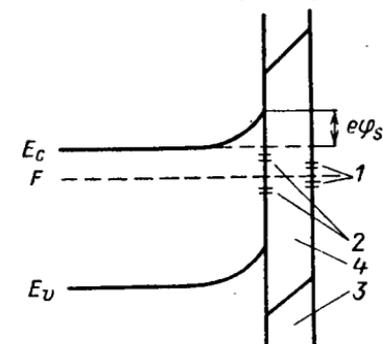


Рис. 9-7. Быстрые и медленные состояния на поверхности германия. 1 — медленные или внешние поверхностные состояния; 2 — быстрые или внутренние поверхностные состояния; 3 — слой окиси германия; 4 — запрещенная зона окиси.

Недостатком измерения поверхностной проводимости в схеме на постоянном токе является неудобство проведения измерений из-за наличия процессов релаксации. От этого недостатка в значительной степени можно избавиться, используя переменное электрическое поле.

9-3. ПОВЕРХНОСТНАЯ РЕКОМБИНАЦИЯ

Рекомбинация носителей заряда в большинстве полупроводников происходит через примесные центры, уровни которых расположены вблизи середины запрещенной зоны. При таком механизме рекомбинации время жизни зависит от концентрации центров рекомбинации. При этом оно тем меньше, чем больше концентрация центров рекомбинации.

На поверхности реальных полупроводников всегда имеется большое количество быстрых и медленных поверхностных состояний, обусловленных адсорбированными примесными атомами или другими дефектами. Центрами рекомбинации могут являться только быстрые уровни, так как время перехода носителей на медленные уровни очень велико. И если концентрация рекомбинационных центров в объеме мала, то быстрые состояния на поверхности будут играть существенную роль в процессах рекомбинации. Рекомбинация через поверхностные уровни, поскольку их количество велико, происходит особенно быстро. В тонких образцах, толщина которых сравнима с длиной диффузионного смещения, имеет место в основном рекомбинация электронов и дырок на поверхности кристалла.

Рассмотрим электронный полупроводник, у которого на поверхности имеются центры рекомбинации одного типа, обменивающиеся носителями с зоной проводимости. Пусть концентрация этих центров N_s , а положение их энергетических уровней таково, как это изображено на рис. 9-8.

Процесс рекомбинации в этом случае происходит следующим образом. Поскольку в приповерхностной области сосредоточен положительный объемный заряд, то электрическое поле направлено от поверхности в глубь полупроводника. Под воздействием этого поля электроны зоны проводимости перемещаются к поверхности кристалла. Одновременно дырки в валентной зоне, стремясь занять квантовые состояния с меньшей энергией, также перемещаются к поверхности. (Это движение электронов и дырок на рис. 9-8 показано пунктирными стрелками.) Затем поверхностный центр рекомбинации захватывает электрон. В акте рекомбинации электрон из

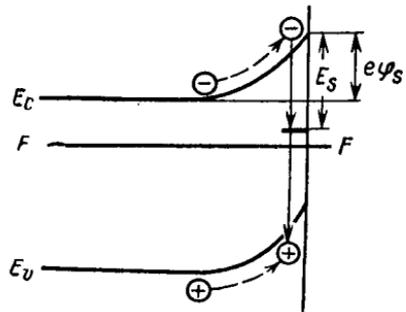


Рис. 9-8. Рекомбинация через поверхностные уровни.

зоны проводимости переходит на энергетический уровень поверхностного центра рекомбинации, а затем с него в валентную зону. Если поверхность кристалла не соприкасается с каким-либо проводником, то полный ток через нее равен нулю. Это означает, что потоки электронов и дырок к поверхности одинаковы и равны числу носителей q_s , рекомбинирующих на 1 см^2 поверхности в 1 сек . При проведении подсчета времени жизни для рекомбинации необходимо также учитывать и обратные потоки электронов и дырок, идущих с поверхностных уровней в объем полупроводника. Этот расчет фактически будет таким же, как и для объемной рекомбинации (см. § 6-7).

В случае поверхностной рекомбинации, так же как и для объемной [выражение (6-15)], если концентрации электронов и дырок у самой поверхности n_s и p_s отличаются от их равновесных значений n_{0s} и p_{0s} , процессы тепловой генерации не уравнивают процесс рекомбинации и число пар носителей, рекомбинирующих на 1 см^2 поверхности в 1 сек , будет равно:

$$q_s = - \left(\frac{dn_s}{dt} \right)_r = - \left(\frac{dp_s}{dt} \right)_r = \gamma_s (n_s p_s - n_{0s} p_{0s}), \quad (9-32)$$

где γ_s — константа, имеющая размерность $\text{см}^4/\text{сек}$.

Если скорость поверхностной рекомбинации невелика, то число пар носителей, рекомбинирующих на поверхности, невелико по сравнению с потоками к поверхности и от поверхности. А эти потоки связаны с дрейфовыми и диффузионными токами в области пространственного заряда. Поэтому при малой скорости поверхностной рекомбинации можно считать, что диффузионные и дрейфовые токи практически уравнивают друг друга. В этом случае концентрации носителей будут определяться распределением Больцмана (9-4). И если величина изгиба зоны у поверхности $e\varphi_s$, а концентрация дырок и электронов в объеме n и p , то их поверхностные значения на основании формулы (9-4) будут равны:

$$n_s = n e^{-\frac{e\varphi_s}{k_0 T}}; \quad p_s = p e^{\frac{e\varphi_s}{k_0 T}}. \quad (9-33)$$

Соответственно для равновесных концентраций

$$n_{0s} = n_0 e^{-\frac{e\varphi_s}{k_0 T}}; \quad p_{0s} = p_0 e^{\frac{e\varphi_s}{k_0 T}}. \quad (9-34)$$

Учитывая выражения (9-33) и (9-34), для равенства (9-32) получим:

$$q_s = \gamma_s (np - n_0 p_0). \quad (9-35)$$

Из равенства (9-35) следует, что поверхностная рекомбинация определяется рекомбинационными явлениями в объеме полупроводника. В случае отсутствия центров прилипания можно считать, что $\Delta n = \Delta p$. Тогда выражение (9-35) примет вид:

$$q_s = \gamma_s (n_0 + p_0 + \Delta n) \Delta n = s \Delta n, \quad (9-36)$$

где

$$s = \gamma_s (n_0 + p_0 + \Delta n) \quad (9-37)$$

имеет размерность скорости и называется скоростью поверхностной рекомбинации.

Число рекомбинирующих на поверхности носителей заряда подобно выражению (6-80) определяется следующим соотношением:

$$q_s = \frac{\gamma_{ns} \gamma_{ps} N_s (n_s p_s - n_{0s} p_{0s})}{\gamma_{ns} (n_s + n_1) + \gamma_{ps} (p_s + p_1)}, \quad (9-38)$$

где γ_{ns} , γ_{ps} — вероятности захвата уровнем электронов и дырок из соответствующих зон в течение одной секунды, имеющие размерность $см^2/сек$; выражения

$$n_1 = N_c e^{-\frac{E_s}{k_0 T}} \quad \text{и} \quad p_1 = N_v e^{-\frac{\Delta E_g - E_s}{k_0 T}}$$

представляют собой равновесные концентрации электронов и дырок, когда уровень Ферми совпадает с уровнем центра рекомбинации E_s ; ΔE_g — ширина запрещенной зоны.

С учетом равенства (9-33) уравнение (9-38) примет вид:

$$q_s = \frac{\gamma_{ns} \gamma_{ps} N_s (np - n_0 p_0)}{\gamma_{ns} n_1 + \gamma_{ps} p_1 + \gamma_{ns} n e^{-\frac{e\varphi_s}{k_0 T}} + \gamma_{ps} p e^{\frac{e\varphi_s}{k_0 T}}}. \quad (9-39)$$

Отсюда найдем скорость поверхностной рекомбинации, которая будет равна:

$$s = \frac{\gamma_{ns} \gamma_{ps} N_s (n_0 + p_0 + \Delta n)}{\gamma_{ns} n_1 + \gamma_{ps} p_1 + \gamma_{ns} n e^{-\frac{e\varphi_s}{k_0 T}} + \gamma_{ps} p e^{\frac{e\varphi_s}{k_0 T}}}. \quad (9-40)$$

При малой концентрации избыточных носителей, т. е. когда $\Delta n \ll n_0 + p_0$, скорость поверхностной рекомбинации не зависит от Δn .

Таким образом, скорость поверхностной рекомбинации зависит от величины поверхностного потенциала φ_s . В том случае, если $\gamma_{ns} n = \gamma_{ps} p$, величина s достигает максимума. На рис. 9-9 изображена зависимость скорости поверхностной рекомбинации от поверхностного потенциала для различных значений энергетического положения рекомбинационного уровня.

Из рис. 9-9 видно, что ширина кривой увеличивается при удалении уровня центра рекомбинации от середины запрещенной зоны.

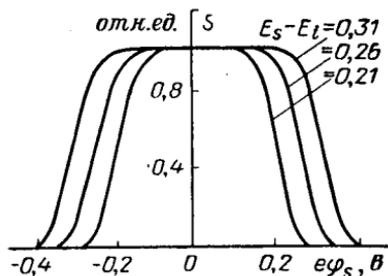


Рис. 9-9. Зависимость скорости поверхностной рекомбинации от поверхностного потенциала для различных значений $E_s - E_i$ (E_i — энергия, соответствующая середине запрещенной зоны).

9-4. ВЛИЯНИЕ ПОВЕРХНОСТНОЙ РЕКОМБИНАЦИИ НА РАСПРЕДЕЛЕНИЕ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА

Рассмотрим плоскую поверхность полупроводника, и пусть избыточные носители генерируются равномерно по всему объему образца со скоростью g . Запишем уравнение непрерывности для некоторого объема полупроводника, находящегося на расстоянии x от поверхности. В этом случае

$$\frac{\partial n}{\partial t} = \frac{1}{e} \frac{\partial j_n}{\partial x} - \frac{n - n_0}{\tau_n} + g. \quad (9-41)$$

Решение проведем для стационарного случая, когда $\frac{\partial n}{\partial t} = 0$. Кроме того, поскольку напряженность электрического поля в области пространственного заряда невелика, дрейфовую составляющую тока можно не учитывать. Тогда $j_n = eD_n \frac{dn}{dx}$.

Соответственно уравнение (9-41) теперь запишется в виде

$$D_n \frac{d^2 \Delta n}{dx^2} - \frac{\Delta n}{\tau_n} + g = 0, \quad (9-42)$$

где $\Delta n = n - n_0$ — избыточная концентрация электронов.

Решение, удовлетворяющее уравнению (9-42), будет иметь вид:

$$\Delta n = \Delta n_1 e^{-\frac{x}{L_n}} + g\tau_n. \quad (9-43)$$

Для нахождения Δn_1 воспользуемся граничными условиями. Будем считать, что избыточные носители, подходя к поверхности, полностью рекомбинируют. Это значит, что

$$D_n \frac{d \Delta n}{dx} = \frac{j_n}{e} = q_s = s \Delta n. \quad (9-44)$$

Из уравнения (9-43) с учетом выражения (9-44) при $x = 0$ получим:

$$\frac{-D_n \Delta n_1}{L_n} = s(\Delta n_1 + g\tau_n), \quad (9-45)$$

откуда найдем Δn_1 , учитывая, что $D_n \tau_n = L_n^2$:

$$\Delta n_1 = -\frac{sg\tau_n L_n}{D_n + sL_n} = -\frac{sg\tau_n^2}{L_n + s\tau_n}. \quad (9-46)$$

Подставив выражение (9-46) в уравнение (9-43), имеем:

$$\Delta n = g\tau_n \left[1 - \frac{s\tau_n}{L_n + s\tau_n} e^{-\frac{x}{L_n}} \right]. \quad (9-47)$$

Это уравнение можно записать через скорость движения электронов v_n , которая равна L_n/τ_n . В этом случае Δn будет определяться

выражением

$$\Delta n = g\tau_n \left[1 - \frac{s}{v_n} e^{-\frac{x}{L_n}} \right]. \quad (9-48)$$

Концентрацию избыточных электронов на поверхности полупроводника Δn_s можно определить, если в выражении (9-47) или в выражении (9-48) положить $x = 0$. Она будет равна:

$$\Delta n_s = \frac{g\tau_n L_n}{L_n + s\tau_n} = \frac{g}{1 + \frac{s}{v_n}} \tau_n. \quad (9-49)$$

Из соотношения (9-49) следует, что поверхностная рекомбинация влияет на процесс восстановления равновесия. Уменьшение величины избыточных электронов вблизи поверхности, обусловленное поверхностной рекомбинацией, заметно сказывается только на расстояниях нескольких L_n от поверхности образца. Кроме того, уменьшение Δn зависит также от отношения s/v_n . И если $s \gg v_n$, то Δn_s очень мал и спад концентрации избыточных электронов вблизи поверхности будет значительным, определяясь выражением вида

$$\Delta n = g\tau_n \left[1 - e^{-\frac{x}{L_n}} \right]. \quad (9-50)$$

Полупроводник, у которого наряду с объемной рекомбинацией наблюдается и поверхностная, характеризуется эффективным временем жизни носителей заряда τ_f . Так как эффекты поверхностной и объемной рекомбинации аддитивны, то для них

$$\frac{1}{\tau_f} = \frac{1}{\tau} + v_s, \quad (9-51)$$

где $1/\tau$ и v_s — объемная и поверхностная постоянные спада соответственно.

Один из известных методов определения скорости поверхностной рекомбинации заключается в следующем. На тонком образце производится измерение τ_f . Затем измеряется τ на толстом образце, толщина которого $d \gg L$. Затем из соотношения (9-51) рассчитывается поверхностная постоянная спада v_s , из которой и определяется s . В том случае, если $s \ll 2D/d$, можно пользоваться упрощенной формулой вида

$$\frac{1}{\tau_f} = \frac{1}{\tau} + \frac{2s}{d}. \quad (9-52)$$

Как показывает эксперимент, скорость поверхностной рекомбинации германия меняется в пределах от 10^8 до 10^2 см/сек. Такие высокие значения скорости поверхностной рекомбинации имеют

место у образцов, поверхность которых обработана шлифованием или подвергнута пескоструйной обработке. Образцы, обработанные подходящим травителем, как правило, характеризуются малыми величинами s .

Глава десятая

ОПТИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

10-1. СПЕКТР ПОГЛОЩЕНИЯ СВЕТА

Когда пучок монохроматического излучения проходит через вещество, то вследствие отражения и поглощения его интенсивность уменьшается. Пусть доля энергии, отраженной на границе тела, составляет величину R , которая носит название коэффициента отражения. Если интенсивность падающего света есть I_0 , а отраженного I_R , то

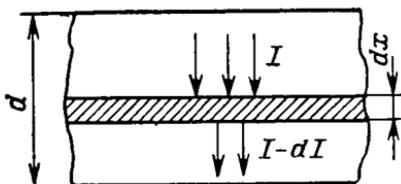


Рис. 10-1. Поглощение света в полупроводнике.

$$R = \frac{I_R}{I_0}. \quad (10-1)$$

Зависимость коэффициента отражения от частоты $R(\omega)$ или длины волны $R(\lambda)$ называется спектром отражения.

Обозначим интенсивность света, падающего на слой dx , как изображено на рис. 10-1, через I . Тогда вследствие поглощения света в этом слое интенсивность излучения уменьшится на величину dI . Количество поглощенной энергии dI пропорционально количеству падающей на слой энергии и толщине поглощенного слоя:

$$-dI = \alpha I dx. \quad (10-2)$$

Коэффициент пропорциональности α , выражающий собой количество поглощенной энергии из пучка единичной интенсивности в слое единичной толщины, носит название коэффициента поглощения.

Проинтегрировав уравнение (10-2)

$$\int_{I_0 - RI_0}^I \frac{dI}{I} = - \int_0^d \alpha dx, \quad (10-3)$$

получим выражение

$$I = I_0 (1 - R) e^{-\alpha d}, \quad (10-4)$$

известное под названием закона Бугера — Ламберта. Величина α является характеристикой поглощающей среды и зависит от длины волны излучения. *Зависимость коэффициента погло-*

чения от частоты $\alpha(\omega)$ или длины волны $\alpha(\lambda)$ называется спектром поглощения вещества.

Пусть имеется N поглощающих центров. Обозначим через σ вероятность поглощения однофотонного пучка на одном поглощающем центре. Другими словами, σ есть эффективное сечение поглощения одного фотона в единицу времени. Эффективное сечение σ зависит от энергии фотона и природы поглощающих центров. Согласно формуле (4-48) $(\sigma N)^{-1}$ является средней длиной свободного пробега фотона l_{Φ} в поглощающей среде, т. е.

$$l_{\Phi} = (\sigma N)^{-1}, \quad (10-5)$$

а коэффициент поглощения

$$\alpha = \frac{1}{l_{\Phi}} \quad (10-6)$$

есть вероятность поглощения фотона на единичной длине. Пусть в полупроводнике имеются поглощающие центры разной природы. Если N_i центры поглощения характеризуются эффективным сечением σ_i , то

$$\alpha_i(\omega) = \sigma_i N_i. \quad (10-7)$$

Полный коэффициент поглощения вещества α есть сумма парциальных коэффициентов поглощения:

$$\alpha(\omega) = \sum_i \alpha_i(\omega) = \sum_i \sigma_i(\omega) N_i. \quad (10-8)$$

Следовательно, *полный спектр поглощения 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 $\hbar\omega$ и импульс $\hbar\vec{\eta}$, обладал энергией E и квазиимпульсом \mathbf{p} , а после взаимодействия имел E' и \mathbf{p} , то эти законы запишутся в виде

$$E' = E + \hbar\omega; \quad (10-9)$$

$$\mathbf{p}' = \mathbf{p} + \hbar\vec{\eta}. \quad (10-10)$$

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 поглощение света решеткой.*

10-2. СОБСТВЕННОЕ ПОГЛОЩЕНИЕ СВЕТА ПРИ ПРЯМЫХ ПЕРЕХОДАХ

Если при поглощении полупроводником кванта излучения электроны валентной зоны, приобретая дополнительную энергию, превышающую или равную ширине запрещенной зоны, переходят в зону

проводимости, такое поглощение называется 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 делятся на два основных вида. У первого из них минимум энергии в зоне проводимости, характеризуемый волновым вектором $\kappa_{\text{мин}}$, и максимум энергии в валентной зоне, определяемый волновым вектором $\kappa_{\text{макс}}$, расположены в одной и той же точке зоны Бриллюэна (обычно в точке $\kappa = 0$). Другими словами, у этих полупроводников $\kappa_{\text{мин}} = \kappa_{\text{макс}}$ (рис. 10-2, а). В качестве примера полупроводника с таким построением зон можно указать антимонид индия (см. рис. 2-20).

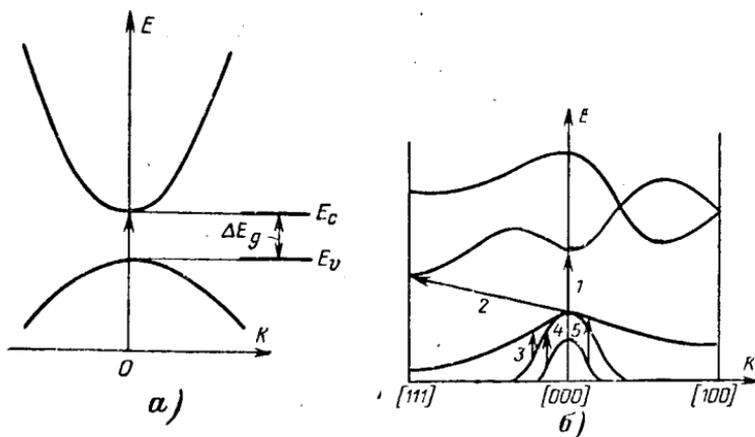


Рис. 10-2. Собственное поглощение при прямых межзонных переходах (а) и оптические переходы в полупроводнике со сложными энергетическими зонами (б).

У второго вида веществ экстремумы зоны проводимости и валентной зоны находятся при различных κ , так что в этом случае $\kappa_{\text{мин}} \neq \kappa_{\text{макс}}$ (рис. 10-2, б). К последнему типу веществ относится большинство полупроводников, в том числе германий (см. рис. 2-16, а) и кремний (см. рис. 2-16, б).

Переходы электронов через запрещенную зону будут происходить прежде всего между энергетическими состояниями, соответствующими максимуму валентной зоны и минимуму зоны проводимости, т. е. при значениях квазиимпульсов ($\mathbf{p} = \hbar\mathbf{k}$) или волнового вектора κ , близких к нулю, как это представлено на рис. 10-2, а.

Волновой вектор фотона $\eta = 2\pi/\lambda_\phi$ очень мал по сравнению с волновым числом электрона, так как длина волны электрона в кристалле при энергии электрона, соответствующей 300°K , составляет примерно $5 \cdot 10^{-7}\text{ см}$, в то время как λ_ϕ — порядка $10^{-1} - 10^{-5}\text{ см}$. В связи с этим в уравнении (10-10) волновым вектором фотона можно пренебречь. Поэтому

$$\kappa' = \kappa \quad (10-11)$$

или

$$\Delta k = k' - k = 0. \quad (10-12)$$

Соотношение (10-11) или (10-12), называемое правилом отбора для электронных переходов, показывает, что в процессе взаимодействия электрона полупроводника с полем излучения возможны только такие переходы, при которых волновой вектор электрона сохраняется. Эти переходы получили название вертикальных или прямых переходов (переход 1 на рис. 10-2, б).

Рассмотрим собственное поглощение при прямых межзонных переходах для полупроводника, у которого зона проводимости и валентная зона обладают сферической симметрией.

Пусть имеют место переходы электронов из состояния в валентной зоне с волновым вектором, принадлежащим интервалу от k до $k + dk$ (рис. 10-3).

В этом случае энергии поглощенного фотона $h\nu$ может быть определена на основании закона сохранения энергии из соотношения

$$\begin{aligned} h\nu &= E_c(k) - E_v(k) = E_c(0) + \frac{\hbar^2 k^2}{2m_n^*} - E_v(0) + \frac{\hbar^2 k^2}{2m_p^*} = \\ &= \Delta E_g + \frac{\hbar^2 k^2}{2m_r^*}, \end{aligned} \quad (10-13)$$

где $\Delta E_g = E_c(0) - E_v(0)$ — ширина запрещенной зоны; $m_r^* = \frac{m_n^* m_p^*}{m_n^* + m_p^*}$ — приведенная эффективная масса электрона и дырки.

Коэффициент поглощения α согласно формуле (10-6) обратно пропорционален длине свободного пробега l_ϕ , величина которой в свою очередь определяется скоростью движения v_ϕ и временем свободного пробега фотона τ_ϕ , т. е.

$$\alpha = \frac{1}{l_\phi} = \frac{1}{\tau_\phi v_\phi} = g(\nu) \frac{\bar{n}}{c}, \quad (10-14)$$

где $g(\nu) = 1/\tau_\phi$ — вероятность поглощения фотона в единицу времени; $v_\phi = c/\bar{n}$ — скорость движения фотона в веществе; \bar{n} — показатель преломления вещества; c — скорость света.

Вероятность поглощения кванта света с энергией, лежащей в интервале от $h\nu$ до $h(\nu + d\nu)$ (рис. 10-3), пропорциональна вероятности перехода электрона $P(\nu)$ и числу квантовых состояний в валентной зоне $N(E')$ в интервале энергий от $-\Delta E_g - E'$ до $-\Delta E_g - E' - dE'$, т. е.

$$g(\nu) d\nu = 2P(\nu) N(E') dE'. \quad (10-15)$$

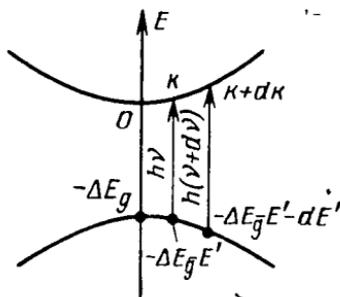


Рис. 10-3. Прямые переходы.

Здесь коэффициент 2 учитывает два возможных направления поляризации света. Плотность квантовых состояний в валентной зоне согласно формуле (3-11) определяется соотношением

$$N(E') = \frac{4\pi (2m_p^*)^{3/2}}{h^3} E'^{1/2}. \quad (10-16)$$

Следовательно,

$$g(\nu) d\nu = \frac{8\pi}{h^3} (2m_p^*)^{3/2} P(\nu) \cdot E'^{1/2} dE' = \frac{2}{\pi^2} P(\nu) \kappa^2 d\kappa, \quad (10-17)$$

откуда, учитывая выражение (10-13), получаем:

$$g(\nu) = \frac{2}{\pi^2} P(\nu) \kappa^2 \frac{d\kappa}{d\nu} = \frac{2(2m_r^*)^{3/2}}{\pi h^2} P(\nu) (h\nu - \Delta E_g)^{1/2}. \quad (10-18)$$

Для разрешенных прямых переходов при $\kappa = 0$ вероятность перехода $P(\nu)$ практически постоянна, поэтому зависимость коэффициента поглощения α_d от энергии у края собственной абсорбции полупроводника на основании выражений (10-14) и (10-18) можно представить в виде.

$$\alpha_d = g(\nu) \frac{\bar{n}}{c} = \frac{2\bar{n} (2m_r^*)^{3/2}}{c\pi h^2} P(\nu) (h\nu - \Delta E_g)^{1/2} = B (h\nu - \Delta E_g)^{1/2}, \quad (10-19)$$

где $B = \frac{2\bar{n} (2m_r^*)^{3/2}}{c\pi h^2} P(\nu)$ — постоянное число.

Соотношение (10-19) справедливо в ограниченном интервале изменения $(h\nu - \Delta E_g)$. Как это видно из рис. 10-4, α_d^2 линейно зависит от $h\nu$ в некотором интервале. Продолжение этой прямой до пересечения с осью абсцисс позволяет определить ширину запрещенной зоны ΔE_g для прямых разрешенных переходов.

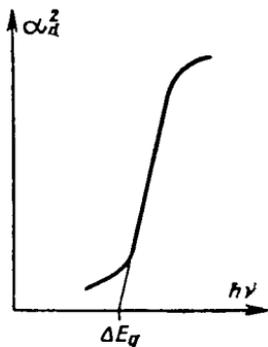


Рис. 10-4. Зависимость α_d^2 от $h\nu$.

Разрешенные вертикальные оптические переходы будут в том случае, если валентная зона полупроводника образуется из s -состояния индивидуальных атомов, а зона проводимости — из p -состояния.

В случае же образования зоны проводимости на основе атомных d -функций переходы будут прямыми, но «запрещенными». При запрещенных переходах для всех $\kappa \neq 0$ вероятность перехода $P(\nu)$ пропорциональна κ^2 , т. е.

$$P(\nu) \sim \kappa^2 = \text{const} (h\nu - \Delta E_g). \quad (10-20)$$

Поэтому коэффициент поглощения для прямых запрещенных межзонных переходов будет зависеть от $h\nu$ согласно соотношению

$$\alpha'_d = C (h\nu - \Delta E_g)^{3/2}, \quad (10-21)$$

где C — постоянная.

На рис. 10-5 изображен спектр поглощения очень чистого монокристалла антимонида индия, измеренный при разных температурах. Для него имеет место крутой рост поглощения, обусловленный прямыми разрешенными переходами.

Однако для антимонида индия положение края полосы поглощения зависит от концентрации примеси. Для материала n -типа при концентрации электронов $5 \cdot 10^{18} \text{ см}^{-3}$ край поглощения лежит в коротковолновой области при $3,5 \text{ мкм}$, а у собственного полупроводника длинноволновая граница при комнатной температуре соответствует $7,1 \text{ мкм}$ (рис. 10-6). Это происходит потому, что в антимониде индия сильно отличаются эффективные массы электрона ($m_n^* = 0,03m$) и дырки ($m_p^* = 0,2m$), а следовательно, и плотности состояний в зоне проводимости и валентной зоне.

На рис. 10-7, кроме зоны проводимости и валентной зоны антимонида индия, изображены пунктиром также зоны, соответствующие полупроводнику, у которого $m_n^* = m_p^* = m$. Поскольку у антимонида индия зона проводимости характеризуется большой крутизной, плотность состояний в ней мала. Вследствие этого даже небольшое количество свободных электронов будет заполнять энергетические уровни зоны проводимости, находящиеся на сравнительно большом расстоянии от ее дна. Это приведет к тому, что с ро-

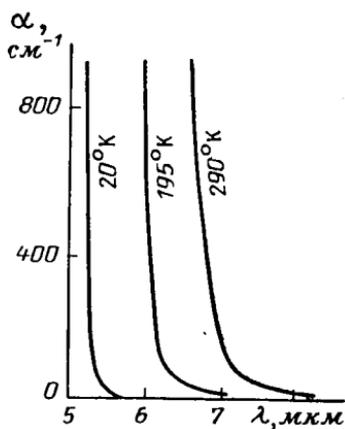


Рис. 10-5. Спектр поглощения антимонида индия при разных температурах.

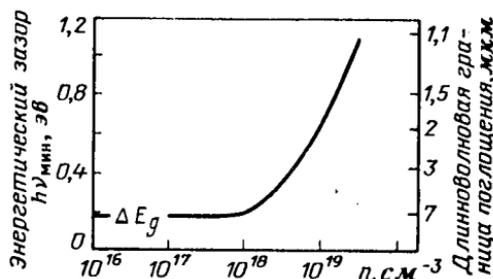


Рис. 10-6. Зависимости оптического энергетического зазора $h\nu_{\text{мин}}$ от концентрации электронов в $\text{InSb } n$ -типа.

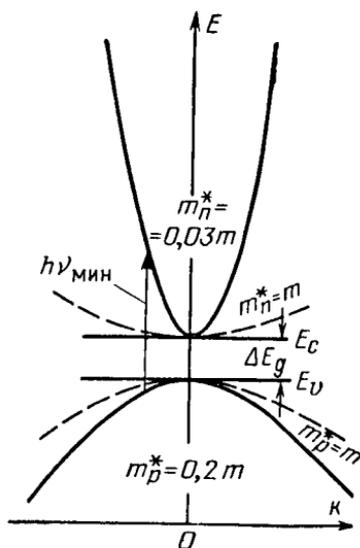


Рис. 10-7. Зависимость энергии от квазиимпульса в антимониде индия.

стом концентрации донорной примеси край собственного поглощения антимионида индия будет смещаться в сторону коротких длин волн. В этом случае величина $h\nu_{\text{мин}}$ (прямой переход на рис. 10-7) будет больше ширины запрещенной зоны.

10-3. СОБСТВЕННОЕ ПОГЛОЩЕНИЕ СВЕТА ПРИ НЕПРЯМЫХ ПЕРЕХОДАХ

В некоторых случаях существуют факторы, «смягчающие» правила отбора, в результате чего оказываются допустимыми также не вертикальные переходы, но вероятность таких переходов намного меньше вероятности прямых переходов (переход 2 на рис. 10-2,б). В отличие от прямых переходов они происходят без сохранения квазиимпульса электрона. Закон сохранения импульса здесь обеспечивается взаимодействием электрона в процессе перехода не только с полем излучения, но и с колебаниями решетки. Другими словами, *непрямые или не-*

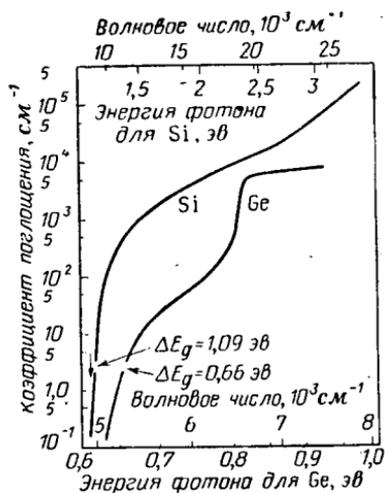


Рис. 10-8. Спектр поглощения германия и кремния.

вертикальные переходы осуществляются с испусканием или поглощением фотона. Эти переходы определяют поглощение, расположенное с длинноволновой стороны у границы собственного поглощения, обусловленного прямыми переходами. Прямые и непрямые переходы достаточно четко проявляются в спектре поглощения германия и кремния. Согласно рис. 10-8 по-

роговые значения энергии фотонов в германии, соответствующие непрямым и прямым оптическим переходам, при комнатной температуре равны 0,62 и 0,81 эВ. При температуре жидкого азота они соответственно составляют 0,72 и 0,88 эВ. При расчете коэффициента поглощения для непрямых переходов, которые сопровождаются поглощением или испусканием одного фотона, 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 сохранения энергии имеет место только для всего перехода в целом, в то время как импульс сохраняется при переходе в виртуальное состояние.

Для простоты положим, что непрямой переход происходит из начального состояния $\kappa_1 = 0$ в конечное с $\kappa_2 = \kappa_{\text{мин}}$ (переход 2' на рис. 10-9). Для такого перехода, идущего с испусканием или

поглощением фонона, возможны два пути. Электрон, находящийся вблизи потолка валентной зоны с $\kappa_1 = 0$, возбуждаясь светом, переходит в зону проводимости, занимая там состояние с тем же волновым вектором (переход 1 на рис. 10-9). В валентной зоне в результате такого перехода остается дырка с $\kappa_1 = 0$. Однако электрон проводимости с $\kappa_1 = 0$ обладает большей энергией, чем энергия, которая соответствует дну зоны проводимости, поэтому он через очень малый промежуток времени совершит переход в состояние вблизи минимума энергии $\kappa_2 = \kappa_{\text{мин}}$, поглощая или испуская при этом фонон с волновым вектором $\kappa_{\text{мин}}$ (переход 2 на рис. 10-9). Кроме того, возможен и такой переход. Поглощая квант света, электрон из глубоко лежащего состояния в валентной зоне вертикально переходит в состояние зоны проводимости с $\kappa_2 = \kappa_{\text{мин}}$ (переход 3 на рис. 10-9). В

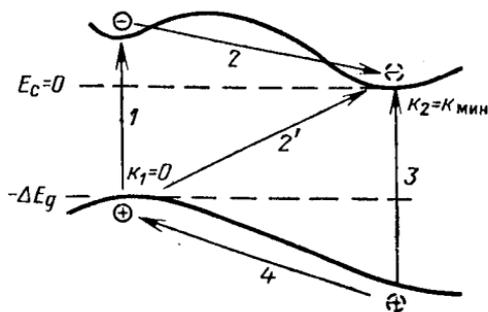


Рис. 10-9. Непрямые оптические переходы электрона из валентной зоны в зону проводимости.

результате в валентной зоне остается глубоко лежащая дырка, которая совершит переход в состояние $\kappa_1 = 0$ вблизи потолка валентной зоны, испуская или поглощая при этом фонон (переход 4 на рис. 10-9). Если энергию фонона обозначить через E_p , то минимальная энергия кванта света, необходимая электрону для совершения непрямого перехода (переход 2' на рис. 10-9) в случае поглощения фонона будет определяться равенством

$$h\nu = \Delta E_g - E_p, \quad (10-22)$$

а в случае испускания фонона

$$h\nu = \Delta E_g + E_p. \quad (10-23)$$

Следовательно, поглощение кванта света, идущее с поглощением фонона, возможно в том случае, если

$$h\nu > \Delta E_g - E_p, \quad (10-24)$$

а излучение фонона будет иметь место только при

$$h\nu > \Delta E_g + E_p. \quad (10-25)$$

Коэффициент поглощения для не прямых переходов α состоит из двух слагаемых, связанных с поглощением (α_a) и испусканием (α_e) фонона:

$$\alpha = \alpha_a + \alpha_e. \quad (10-26)$$

Если учесть соотношения (10-24) и (10-25), то

$$\alpha_a = 0 \text{ при } hv < \Delta E_g - E_p; \quad (10-27)$$

$$\alpha_e = 0 \text{ при } hv < \Delta E_g + E_p. \quad (10-28)$$

Произведем расчет коэффициента поглощения для переходов электрона по второму варианту, т. е. когда вначале совершается вертикальный переход 3, а затем переход 4. Пусть начальная энергия электрона в валентной зоне будет $-\Delta E_g - E'$ (начало отсчета энергии производится от дна зоны проводимости), а E — энергия электронов в конечном состоянии в зоне проводимости. Закон сохранения энергии для этого случая с учетом поглощения или испускания фонона запишется в виде

$$hv = \Delta E_g + E + E' \mp E_p. \quad (10-29)$$

Будем считать значение E фиксированным. Тогда его максимальное значение E_m согласно рис. 10-9 будет равно:

$$E_m = hv - \Delta E_g \pm E_p. \quad (10-30)$$

Здесь знак плюс соответствует поглощению, а знак минус — испусканию фонона.

Если учесть соотношение (10-29), то число состояний в валентной зоне для участка энергии dE' в соответствии с выражением (10-16) равно:

$$N(E') dE' = a' E'^{1/2} dE' = a' (hv - \Delta E_g - E \mp E_p)^{1/2} dE', \quad (10-31)$$

где a' — постоянная.

Кроме того, из формулы (10-29) для частоты в этом интервале следует соотношение $dE' = h dv$. Число пар состояний $N_c(v) dv$, между которыми осуществляются оптические переходы с частотой поглощаемых фононов в интервале от ν до $\nu + dv$, зависит как от числа состояний в валентной зоне $N(E') dE'$, так и от числа состояний в зоне проводимости $N(E) dE$, т. е.

$$N_c(v) dv = N(E') N(E) dE dE' = aa' E'^{1/2} E^{1/2} dE dE'. \quad (10-32)$$

Определим полное число пар состояний $N(v) dv$. Для этого проинтегрируем выражение (10-32) по E зоны проводимости, в которой заняты состояния с энергией по 0 до E_m . В результате получим:

$$\begin{aligned} N(v) dv &= aa' h dv \int_0^{E_m} (hv - \Delta E_g - E \pm E_p)^{1/2} E^{1/2} dE = aa' h dv \times \\ &\times \int_0^{E_m} (E_m - E)^{1/2} E^{1/2} dE = AE_m^2 dv, \end{aligned} \quad (10-33)$$

где A — постоянная.

Коэффициент поглощения для непрямых переходов с поглощением фонона пропорционален плотности тех состояний, из которых совершаются переходы, т. е. полному числу пар состояний $N(v) dv$,

и относительной вероятности поглощения фонона, т. е. количеству N_p фононов с энергией E_p . Так как согласно выражению (4-159)

$$N_p = \frac{1}{\frac{E_p}{e^{\frac{h\nu_0 T}{k_0 T}} - 1}}, \quad (10-34)$$

то коэффициент поглощения для непрямых переходов с поглощением фонона определится соотношением вида

$$\alpha_a \sim N(\nu) N_p P(\nu) = \frac{B (h\nu - \Delta E_g + E_p)^2}{\frac{E_p}{e^{\frac{h\nu_0 T}{k_0 T}} - 1}} \quad (h\nu > \Delta E_g - E_p); \quad (10-35)$$

$$\alpha_a = 0 \quad (h\nu \leq \Delta E_g - E_p).$$

Здесь B — функция, слабо меняющаяся с частотой.

Из теории колебаний кристаллической решетки известно, что отношение вероятности испускания к вероятности поглощения фонона равно $\frac{N_p + 1}{N_p}$. Поэтому коэффициент поглощения для непрямых переходов с испусканием фонона будет равен:

$$\begin{aligned} \alpha_e &= B (h\nu - \Delta E_g - E_p)^2 (N_p + 1) = \\ &= \frac{B (h\nu - \Delta E_g - E_p)^2}{\frac{E_p}{1 - e^{-\frac{h\nu_0 T}{k_0 T}}}} \quad (h\nu > \Delta E_g + E_p); \end{aligned} \quad (10-36)$$

$$\alpha_e = 0 \quad (h\nu \leq \Delta E_g + E_p).$$

Таким образом, коэффициент поглощения α для непрямых межзонных переходов, если учесть выражения (10-35) и (10-36), имеет вид:

$$\alpha = \alpha_a + \alpha_e = B \left[\frac{(h\nu - \Delta E_g + E_p)^2}{\frac{E_p}{e^{\frac{h\nu_0 T}{k_0 T}} - 1}} + \frac{(h\nu - \Delta E_g - E_p)^2}{\frac{E_p}{1 - e^{-\frac{h\nu_0 T}{k_0 T}}}} \right] \quad (10-37)$$

$$(h\nu > \Delta E_g + E_p).$$

Соотношение (10-37) можно представить в виде

$$\alpha = B \frac{(h\nu - \Delta E_g + E_p)^2}{\frac{E_p}{e^{\frac{h\nu_0 T}{k_0 T}} - 1}}, \quad (\Delta E_g - E_p < h\nu \leq \Delta E_g + E_p); \quad (10-38)$$

$$\alpha = 0 \quad (h\nu \leq \Delta E_g - E_p).$$

График зависимости $\alpha^{1/2}$ от $h\nu$ имеет два прямолинейных участка, как это показано на рис. 10-10.

Прямая 1 соответствует зависимости $\alpha^{1/2}$, которая пересекается с осью абсцисс в точке $h\nu = \Delta E_g + E_p$. Прямая 2 характеризует зависимость $\alpha^{1/2}$ и пересекается с осью абсцисс в точке

$h\nu = \Delta E_g - E_p$. Длина отрезка между точками пересечения прямых $\alpha_a^{1/2}$ и $\alpha_e^{1/2}$, соответствующих одной и той же температуре, с осью $h\nu$ равна $2E_p$. Посередине этого отрезка находится точка $h\nu = \Delta E_g$.

При понижении температуры график зависимости $\alpha^{1/2}$ смещается в сторону больших значений $h\nu$. При этом, как следует из рис. 10-10, наклон прямой $\alpha^{1/2}$ стремится к нулю. Это происходит в силу того, что при низких температурах возбуждается очень мало фононов, а поэтому поглощение фонона маловероятно. Кроме того,

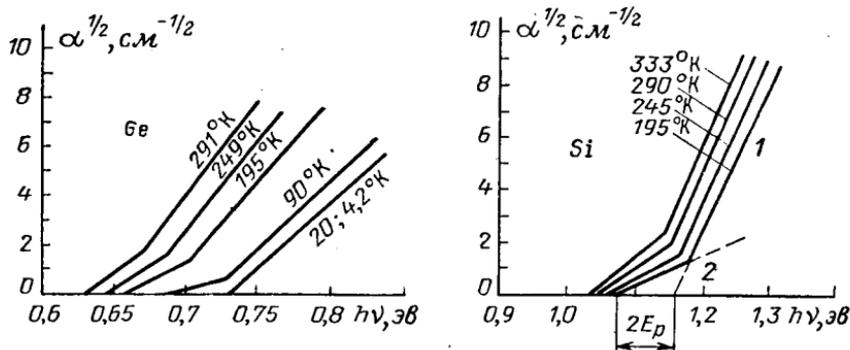


Рис. 10-10. Зависимость $\alpha^{1/2}$ от температуры для германия и кремния.

поскольку ширина запрещенной зоны ΔE_g зависит от температуры, точки пересечения прямых $\alpha^{1/2}$ и $\alpha_e^{1/2}$ с осью $h\nu$ сдвигаются с температурой.

Таблица 10-1

Энергии фононов и дебаевские температуры для различных ветвей колебаний в германии и кремнии

Полупроводник	Тип колебаний	Величина энергии фононов $E_p = k_B \theta$, эв	Эквивалентная температура θ , °K
Германий	Продольные акустические (t_a)	0,0275	320
	Поперечные акустические (t_a)	0,0077	90
Кремний	Поперечные оптические (t_o)	0,034	400
	Продольные акустические (t_a)	0,016	230
	Поперечные акустические (t_a)	0,055	700
	Поперечные оптические (t_o)	0,119	1370
	Продольные оптические (l_o)	0,083	1010

Если прямой переход при $\kappa_1 = 0$ запрещен, то не прямые запрещенные межзонные переходы для этого случая будут характеризоваться коэффициентом поглощения, который согласно выражению

(10-20) будет пропорционален

$$\alpha \sim (h\nu - \Delta E_g \pm E_p)^3. \quad (10-39)$$

Для прямых межзонных переходов коэффициент поглощения обычно достигает величины $10^4 - 10^5 \text{ см}^{-1}$, а для непрямых межзонных переходов он составляет $10^{-1} - 10^3 \text{ см}^{-1}$ (см. рис. 10-8).

Мы рассмотрели один тип фононов. В действительности же кристаллическая решетка полупроводника имеет две ветви колебаний: акустическую и оптическую, каждая из которых обладает двумя поперечными (t) и одной продольной (l) ветвью колебаний. Поэтому при определении коэффициента поглощения для непрямых межзонных переходов необходимо учитывать все ветви колебаний. В табл. 10-1 приведены величины энергии фононов E_p , а также дана эквивалентная им температура $\theta = E_p/k_0$ (температура Дебая) для германия и кремния.

10-4. ЭКСИТОННОЕ ПОГЛОЩЕНИЕ СВЕТА

При поглощении света полупроводником возможно такое возбуждение электрона валентной зоны, при котором он не переходит в зону проводимости, а образует с дыркой связанную систему. Такая система получила название экситона. Если размеры экситона велики по сравнению с постоянной решетки, то взаимодействие электрона и дырки можно представить как кулоновское взаимодействие двух точечных зарядов, ослабленное в $\epsilon\epsilon_0$ раз, где ϵ_0 — диэлектрическая постоянная, ϵ — относительная диэлектрическая проницаемость кристалла.

Пусть \mathbf{r}_n и \mathbf{r}_p — радиусы-векторы, определяющие положения электрона и дырки, а m_n^* и m_p^* — их эффективные массы. Тогда уравнение Шредингера для взаимодействующих электрона и дырки имеет вид:

$$\left[-\frac{\hbar^2}{2m_n^*} \nabla_n^2 - \frac{\hbar^2}{2m_p^*} \nabla_p^2 - \frac{e^2}{4\pi\epsilon\epsilon_0 r} \right] \Psi(\mathbf{r}_n, \mathbf{r}_p) = E_{ex} \Psi(\mathbf{r}_n, \mathbf{r}_p), \quad (10-40)$$

где ∇_n^2 и ∇_p^2 — операторы Лапласа в координатах электрона и дырки; E_{ex} — энергия экситона.

Если ввести радиус-вектор центра тяжести \mathbf{R} , характеризующий положение электронно-дырочной пары, и радиус-вектор положения электрона относительно дырки \mathbf{r} как

$$\mathbf{R} = \frac{m_n^* \mathbf{r}_n + m_p^* \mathbf{r}_p}{m_n^* + m_p^*}, \quad (10-41)$$

$$\mathbf{r} = \mathbf{r}_n + \mathbf{r}_p, \quad (10-42)$$

то уравнение (10-40) в новых координатах будет иметь вид:

$$\left[-\frac{\hbar^2}{2M^*} \nabla_R^2 - \frac{\hbar^2}{2m_r^*} \nabla_r^2 - \frac{e^2}{4\pi\epsilon\epsilon_0 r} \right] \Psi(\mathbf{R}, \mathbf{r}) = E_{ex} \Psi(\mathbf{R}, \mathbf{r}), \quad (10-43)$$

где $M^* = m_n^* + m_p^*$ — эффективная масса экситона; m_r^* — приведенная масса электрона и дырки;

$$\frac{1}{m_r^*} = \frac{1}{m_n^*} + \frac{1}{m_p^*}.$$

Уравнение (10-43) имеет решение

$$\Psi(\mathbf{R}, \mathbf{r}) = e^{i\mathbf{k}_{ex}\cdot\mathbf{R}}\Psi(\mathbf{r}), \quad (10-44)$$

отвечающее свободному (трансляционному) движению центра тяже-

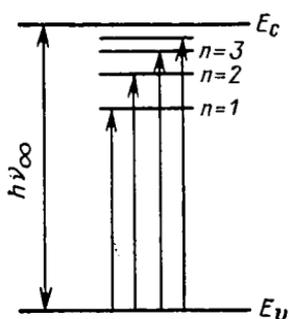


Рис. 10-11. Уровни энергии экситона.

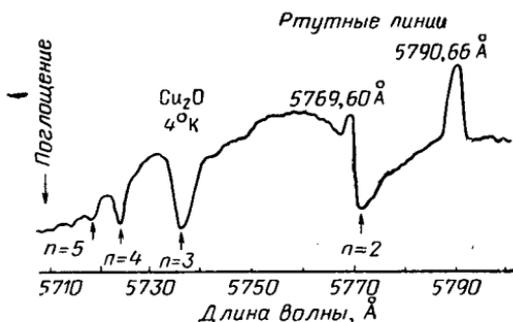


Рис. 10-12. Спектр экситонного поглощения в закиси меди.

сти экситона с волновым вектором \mathbf{k}_{ex} . Для водородоподобной системы $\Psi(\mathbf{r})$ есть решение уравнения

$$\left[-\frac{\hbar}{2m_r^*} \nabla_r^2 - \frac{e^2}{4\pi\epsilon_0 r} \right] \Psi(\mathbf{r}) = E\Psi(\mathbf{r}). \quad (10-45)$$

Уравнение (10-45) описывает внутреннее движение экситона с собственными значениями энергии E , образующими водородоподобную серию:

$$E = -\frac{e^4 m}{32\hbar^2 \pi^2 \epsilon_0^2 e_0^2} \frac{1}{n^2} = -\frac{R'}{n^2}, \quad (10-46)$$

где $R' = \frac{e^4 m}{32\hbar^2 \pi^2 \epsilon_0^2 e_0^2}$ — постоянная; n — главное квантовое число.

Тогда полная энергия экситона

$$E_{ex} = \frac{\hbar^2 \kappa_{ex}^2}{2M^*} - \frac{R'}{n^2}. \quad (10-47)$$

При условии изотропности m_n^* , m_p^* , ϵ и, если оптический переход прямой и совершается при $\kappa = 0$, экситон будет характеризоваться водородоподобной серией линий поглощения (рис. 10-11), удовлетворяющих соотношению

$$h\nu = h\nu_{\infty} - \frac{R'}{n^2}, \quad (10-48)$$

где $h\nu_{\infty}$ совпадает с шириной запрещенной зоны при $\kappa = 0$.

На рис. 10-12 изображен спектр экситонного поглощения в закиси меди.

10-5. ПОГЛОЩЕНИЕ СВЕТА СВОБОДНЫМИ НОСИТЕЛЯМИ ЗАРЯДА

При освещении полупроводника светом соответствующей длины волны электроны зоны проводимости и электроны не полностью заполненной валентной зоны могут переходить внутри зоны с одного уровня на другой. Такие внутризонные переходы носителей заряда происходят с нарушением правила отбора. В силу закона сохранения импульса они осуществляются только тогда, когда наряду с поглощением фотона происходит поглощение или испускание фонона. Поэтому поглощение свободными электронами можно рассматривать методами, которые используются для непрямых переходов. Однако результаты получаются те же, если определение коэффициента поглощения свободными электронами производить на основании соотношения для проводимости в высокочастотном электрическом поле.

Как известно, теория распространения электромагнитных волн в проводящей среде основывается на уравнениях Максвелла, которые для случая, когда отсутствуют сторонние поля и объемные заряды, запишутся в виде

$$\operatorname{rot} \mathbf{H} = \sigma \mathcal{E} + \varepsilon \varepsilon_0 \frac{\partial \mathcal{E}}{\partial t}; \quad (10-49)$$

$$\operatorname{rot} \mathcal{E} = -\kappa \kappa_0 \frac{\partial \mathbf{H}}{\partial t}; \quad (10-50)$$

$$\operatorname{div} \mathbf{H} = 0; \quad (10-51)$$

$$\operatorname{div} \mathcal{E} = 0. \quad (10-52)$$

В этих уравнениях σ — удельная проводимость при данной оптической частоте; κ , κ_0 — магнитная проницаемость, а ε , ε_0 — диэлектрическая постоянная среды и вакуума соответственно.

Из уравнений (10-49) и (10-50) следует:

$$\begin{aligned} \operatorname{rot} \operatorname{rot} \mathcal{E} &= \operatorname{rot} \left[-\kappa \kappa_0 \frac{\partial \mathbf{H}}{\partial t} \right] = -\kappa \kappa_0 \frac{\partial}{\partial t} \operatorname{rot} \mathbf{H} = \\ &= -\kappa \kappa_0 \left(\sigma \frac{\partial \mathcal{E}}{\partial t} + \varepsilon \varepsilon_0 \frac{\partial^2 \mathcal{E}}{\partial t^2} \right). \end{aligned} \quad (10-53)$$

С другой стороны, учитывая формулу (10-52), имеем:

$$\operatorname{rot} \operatorname{rot} \mathcal{E} = [\nabla [\nabla \mathcal{E}]] = \nabla (\nabla \mathcal{E}) - \nabla^2 \mathcal{E} = \operatorname{grad} \operatorname{div} \mathcal{E} - \nabla^2 \mathcal{E} = -\nabla^2 \mathcal{E}. \quad (10-54)$$

Следовательно,

$$\nabla^2 \mathcal{E} - \kappa \kappa_0 \sigma \frac{\partial \mathcal{E}}{\partial t} - \kappa \kappa_0 \varepsilon \varepsilon_0 \frac{\partial^2 \mathcal{E}}{\partial t^2} = 0. \quad (10-55)$$

Пусть электромагнитная волна распространяется в направлении x со скоростью v . Для этого случая решением (10-55) будет:

$$\mathcal{E}_x = \mathcal{E}_0 e^{i\omega \left(t - \frac{x}{v} \right)}, \quad (10-56)$$

при условии, если

$$\frac{1}{v^2} = \kappa \kappa_0 \epsilon \epsilon_0 - \frac{i \kappa \kappa_0 \sigma}{\omega}. \quad (10-57)$$

Но $v = c/\bar{n}^*$, где c — скорость света в вакууме, а \bar{n}^* — комплексный показатель преломления среды. Следовательно,

$$\bar{n}^{*2} = c^2 \kappa \kappa_0 \left(\epsilon \epsilon_0 - \frac{i \sigma}{\omega} \right). \quad (10-58)$$

Если учесть, что для вакуума $\bar{n}^* = 1$, $\kappa = 1$, $\epsilon = 1$ и $\sigma = 0$, то $c^2 \epsilon_0 \kappa_0 = 1$ или $c = (\kappa_0 \epsilon_0)^{-1/2}$. Кроме того, при определении оптических свойств полупроводника можно положить $\kappa = 1$. В соответствии с этим уравнение (10-58) запишется следующим образом:

$$\bar{n}^{*2} = \epsilon - \frac{i \sigma}{\epsilon_0 \omega}. \quad (10-59)$$

Комплексный показатель преломления \bar{n}^* можно записать через показатель поглощения света χ в виде

$$\bar{n}^* = \bar{n} - i\chi. \quad (10-60)$$

Тогда

$$\bar{n}^{*2} = \bar{n}^2 - \chi^2 - i2\bar{n}\chi = \epsilon - i \frac{\sigma}{\epsilon_0 \omega}, \quad (10-61)$$

откуда следует, что

$$\bar{n}^2 = \chi^2 + \epsilon; \quad (10-62)$$

$$\chi = \frac{\sigma}{2\bar{n}\epsilon_0 \omega}. \quad (10-63)$$

Теперь подставим выражение (10-60) в уравнение (10-56). В результате будем иметь:

$$\mathcal{E}_x = \mathcal{E}_0 e^{i\omega \left(t - \frac{\bar{n}^*}{c} x \right)} = A e^{i\omega \left(t - \frac{n}{c} x \right)}, \quad (10-64)$$

где

$$A = \mathcal{E}_0 e^{-\frac{\omega \chi}{c} x} \quad (10-65)$$

есть амплитуда электрического вектора.

Поскольку амплитуда магнитного вектора также определяется величиной $H_0 e^{-\frac{\omega \chi}{c} x}$, то поток энергии волны, который пропорционален амплитуде электрического и магнитного векторов, уменьшается с расстоянием как $e^{-\frac{2\omega \chi}{c} x}$. В то же время интенсивность света, прошедшего через среду толщиной x , пропорциональна $e^{-\alpha x}$. Следовательно, коэффициент поглощения

$$\alpha = \frac{2\omega \chi}{c} = \frac{4\pi \chi}{\lambda_0}, \quad (10-66)$$

где λ_0 — длина волны света в вакууме.

Если учесть формулу (10-63), получим:

$$\alpha = \frac{\sigma}{c\bar{n}\epsilon_0}. \quad (10-67)$$

Таким образом, коэффициент поглощения света определяется проводимостью вещества.

Поскольку проводимость вещества $\sigma = en\mu = \frac{e^2 \langle \tau \rangle}{m^*} n$, то коэффициент поглощения можно записать через концентрацию свободных носителей заряда n в виде

$$\alpha = \frac{\sigma}{c n \epsilon_0} = \frac{e^2 \langle \tau \rangle}{c n \epsilon_0 m^*} n. \quad (10-68)$$

Следовательно, коэффициент поглощения света свободными носителями заряда с концентрацией n зависит от их эффективной массы, среднего времени релаксации и показателя преломления среды \bar{n} . Но показатель преломления зависит от длины волны, поэтому и коэффициент поглощения свободными носителями заряда должен зависеть от длины волны.

Для того чтобы определить зависимость проводимости σ или показателя поглощения χ от длины падающего света, рассмотрим поляризацию вещества. Рассмотрение проведем для случая, когда длина волны падающего излучения очень велика и находится вдали от края полосы его 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

$$m_n^* \frac{d^2 x}{dt^2} + m_n^* g \frac{dx}{dt} = e \mathcal{E} e^{i\omega t}. \quad (10-69)$$

Решение этого уравнения будем искать в виде

$$x = B e^{i\omega t}. \quad (10-70)$$

Подставляя выражение (10-70) в уравнение (10-69), будем иметь:

$$-m_n^* \omega^2 B + i m_n^* g \omega B = e \mathcal{E},$$

откуда найдем, что

$$B = \frac{e \mathcal{E}}{m_n^* \omega (-\omega + i g)}. \quad (10-71)$$

Если в единице объема среды имеется n свободных электронов, то поляризуемость ее составляет $P = \frac{enB}{\mathcal{E}}$, а диэлектрическая постоянная равна $\epsilon^* = \epsilon_0 + P$.

Таким образом,

$$\epsilon^* = \epsilon_0 + \frac{e^2 n}{m_n^* \omega (-\omega + i g)}. \quad (10-72)$$

Но комплексный показатель преломления среды выражается соотношением $\bar{n}^* = (\epsilon^*/\epsilon_0)^{1/2}$, поэтому

$$(\bar{n} - i\chi)^2 = \frac{\epsilon^*}{\epsilon_0} = 1 + \frac{e^2 n}{\epsilon_0 m_n^* \omega (-\omega + ig)}, \quad (10-73)$$

откуда следует, что

$$\bar{n}^2 - \chi^2 - 1 = - \frac{e^2 n}{\epsilon_0 m_n^* (\omega^2 + g^2)}; \quad (10-74)$$

$$2\bar{n}\chi = \frac{e^2 ng}{\epsilon_0 m_n^* \omega (\omega^2 + g^2)}. \quad (10-75)$$

В случае стационарного поля, для которого из (10-69) имеем:

$$m_n^* \frac{dx}{dt} = e\mathcal{E}, \quad (10-76)$$

скорость движения электрона проводимости равна

$$v = dx/dt = \mu_n \mathcal{E},$$

где μ_n — подвижность электрона при данной частоте.

Так как

$$g = \frac{e\mathcal{E}}{m_n^* v} = \frac{e}{m_n^* \mu_n}, \quad (10-77)$$

то соотношение (10-75) можно записать в виде

$$2\bar{n}\chi = \frac{en\mu_n}{\epsilon_0 \omega \left[1 + \left(\frac{m_n^* \mu_n \omega}{e} \right)^2 \right]} = \frac{\sigma}{\epsilon_0 \omega \left[1 + \left(\frac{m_n^* \mu_n \omega}{e} \right)^2 \right]}. \quad (10-78)$$

Соответственно

$$\alpha = \frac{2\omega\chi}{c} = \frac{\sigma}{c\bar{n}\epsilon_0 \left[1 + \left(\frac{m_n^* \mu_n \omega}{e} \right)^2 \right]}. \quad (10-79)$$

Если $\frac{m_n^* \mu_n \omega}{e} \gg 1$ и учитывая, что $\omega = 2\pi\nu = \frac{2\pi c}{\lambda}$, из соотношения (10-79) получаем:

$$\alpha = \frac{e^3 n}{4\pi^2 c^3 \epsilon_0 \bar{n} m_n^* \mu_n} \lambda^2. \quad (10-80)$$

Из соотношения (10-80) следует, что коэффициент поглощения свободными носителями заряда пропорционален их концентрации, квадрату длины волны падающего света и обратно пропорционален подвижности носителей и квадрату их эффективной массы.

Теоретическая зависимость α от λ^2 хорошо согласуется с данными эксперимента. На рис. 10-13 представлена зависимость коэффициента поглощения электронами проводимости в сурьмянистом индии n -типа от λ^2 . Прямая на рис. 10-13, а соответствует концентрации электронов $3,5 \cdot 10^{17} \text{ см}^{-3}$, на рис. 10-13, б — концентрации $6,2 \cdot 10^{17} \text{ см}^{-3}$.

Если же энергетические зоны у полупроводника сложные, как, например, у германия и кремния, то на поглощение свободными носителями накладывается поглощение, обусловленное переходами дырок между отдельными подзонами сложной валентной зоны. Для этих переходов соблюдаются правила отбора. На рис. 10-2, б они изображены стрелками (переходы 3, 4 и 5). В качестве иллюстрации такого поглощения на рис. 10-14 приведена кривая абсорбции для

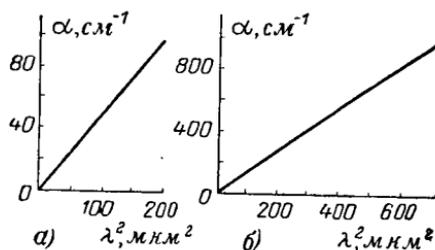


Рис. 10-13. Зависимость поглощения электронами проводимости в сурьмянистом индии *n*-типа от λ^2 для концентрации электронов $3,5 \cdot 10^{17} \text{ см}^{-3}$ (а) и $6,2 \cdot 10^{17} \text{ см}^{-3}$ (б).

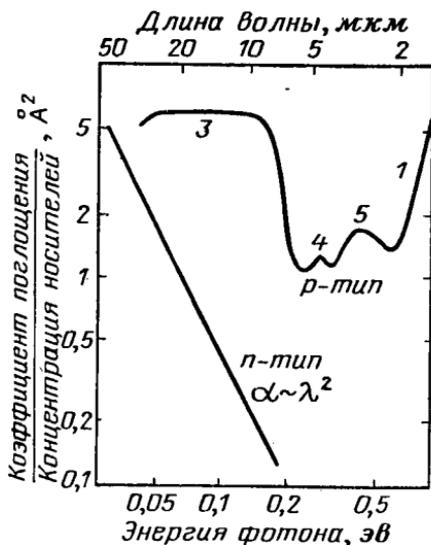


Рис. 10-14. Поглощение свободными носителями заряда в германии *n*- и *p*-типа.

образцов германия *n*- и *p*-типа. На этом рисунке участок 1 является собственным поглощением германия, а максимумы полос абсорбции 3, 4 и 5 определяются переходами дырок между соответствующими энергетическими подзонами валентной зоны.

10-6. ПРИМЕСНОЕ ПОГЛОЩЕНИЕ

Оптическое поглощение, приводящее к ионизации или возбуждению примесных центров в кристалле, называется примесным. Если в кристаллической решетке полупроводника имеется донорная или акцепторная примесь, то при освещении полупроводника светом соответствующей длины волны будет наблюдаться полоса поглощения, вызываемая переходами электронов с энергетических уровней доноров в зону проводимости или переходами электронов из валентной зоны на акцепторные 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 согласно

(2-135) определяется выражением

$$E_d = - \frac{13,52Z^2}{e^2} \left(\frac{m^*}{m} \right) \frac{1}{n^2}, \quad (10-81)$$

где $n = 1, 2, 3 \dots$

Возбуждение электронов примеси, т. е. перевод электрона из основного состояния в возбужденное, приведет к появлению поглощения, в спектре которого будет наблюдаться несколько полос.

В качестве примера примесного поглощения на рис. 10-15 изображены спектры для кремния, легированного различными акцепторными и донорными примесями.

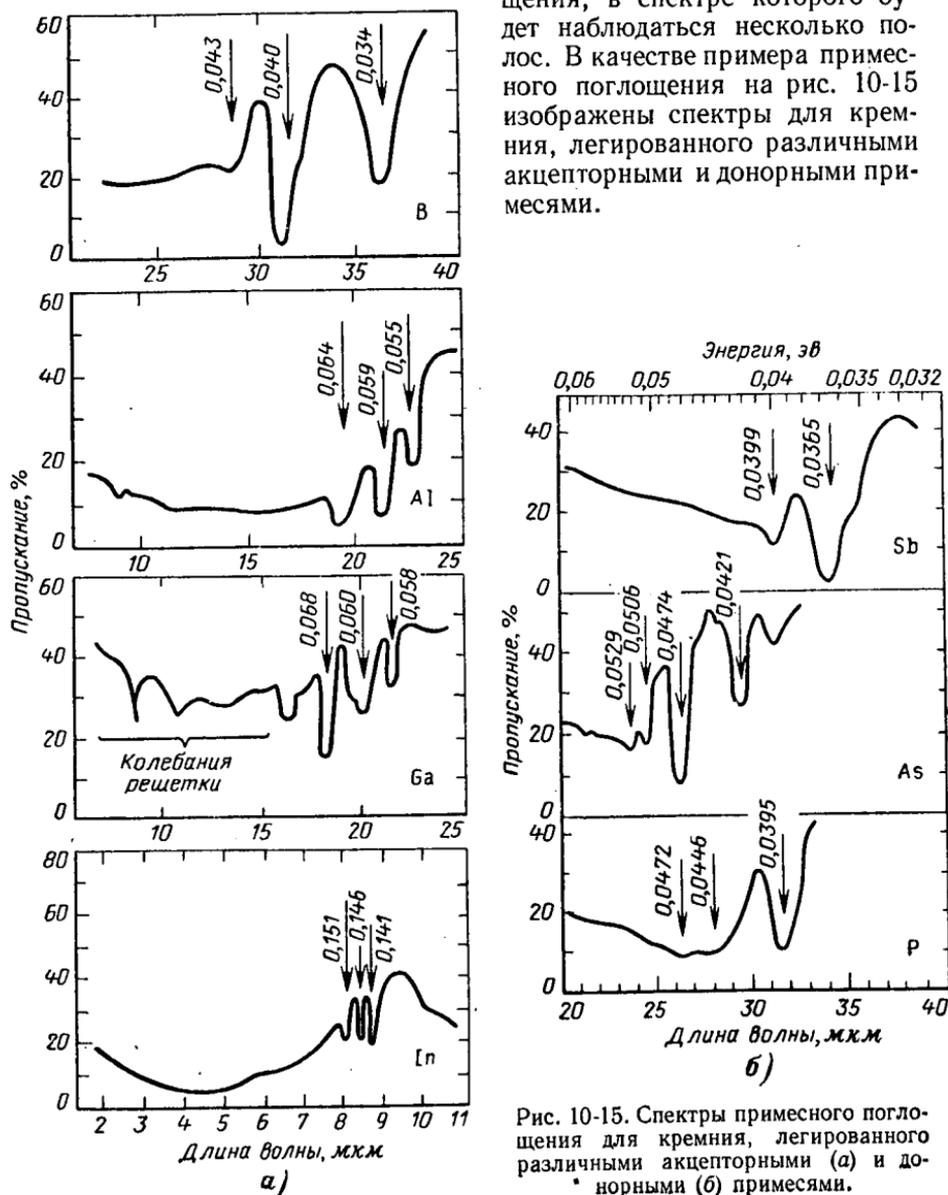


Рис. 10-15. Спектры примесного поглощения для кремния, легированного различными акцепторными (а) и донорными (б) примесями.

10-7. ПОГЛОЩЕНИЕ СВЕТА РЕШЕТКОЙ

Если оптическое поглощение полупроводника связано с изменением колебательной энергии атомов его решетки, то такое поглощение называется поглощением света решеткой. В связи с тем, что решетка может поглощать энергию электромагнитного поля излучения только при определенных значениях энергии фотона, спектр ее характеризуется рядом пиков поглощения, которые обычно накладываются на поглощение свободными носителями. Интенсивность и положение полос поглощения решетки не меняются с изменением концентрации в ней примеси вплоть до величины 10^{18} см^{-3} . Не зависит это поглощение и от концентрации несовершенств кристалла полупроводника, если она не превышает величины 10^{19} см^{-3} . На рис. 10-16 изображены спектры поглощения света решеткой германия и кремния.

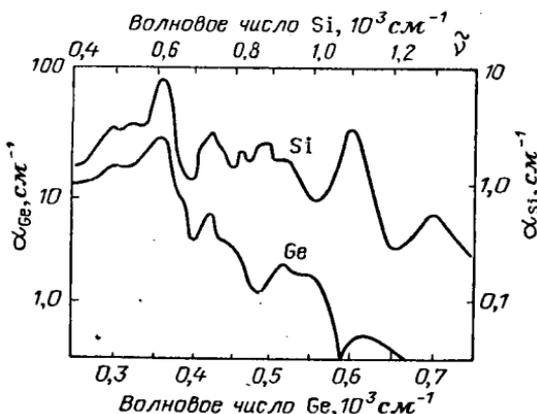


Рис. 10-16. Поглощение решеткой германия и кремния.

10-8. ЗАВИСИМОСТЬ ПОЛОЖЕНИЯ КРАЯ СОБСТВЕННОГО ПОГЛОЩЕНИЯ ОТ ТЕМПЕРАТУРЫ И ДАВЛЕНИЯ

Многочисленные опыты, проведенные на разных полупроводниках, показывают, что температура и давление оказывают существенное влияние на спектр собственного поглощения. Знание влияния температуры и давления на поглощение полупроводников необходимо как для решения теоретических задач, связанных с рассмотрением поведения электронов в электрическом поле кристаллической решетки, так и для решения задач практического значения, связанных с изготовлением и эксплуатацией полупроводниковых приборов. Влияние температуры и давления на их физические параметры обусловлено изменением ширины запрещенной зоны под действием этих факторов.

Температурная зависимость смещения края полосы поглощения определяется тем, что при повышении температуры происходит расширение кристалла, а также имеет место изменение взаимодействия электронов с решеткой. Как оказывается, при нагревании веществ энергетические зоны, как правило, расширяются, что и приводит к уменьшению ширины запрещенной зоны, а это вызывает

смещение полосы фундаментального поглощения в сторону длинных волн (см. рис. 10-8). Для некоторой области температур изменение ширины запрещенной зоны имеет линейный характер и подчиняется соотношению

$$\Delta E_g = \Delta E_{g0} - \alpha T, \quad (10-82)$$

где ΔE_{g0} — ширина запрещенной зоны при 0° К . Для кремния $\alpha = 4 \cdot 10^{-4} \text{ эв/град}$, а для германия $\alpha = 4,5 \cdot 10^{-4} \text{ эв/град}$.

Ширина запрещенной зоны с изменением давления меняется как

$$\Delta E_g = \Delta E_g|_{P=0} + \gamma P. \quad (10-83)$$

Поэтому у большинства полупроводников имеет место смещение полосы собственного поглощения при увеличении давления в сторону

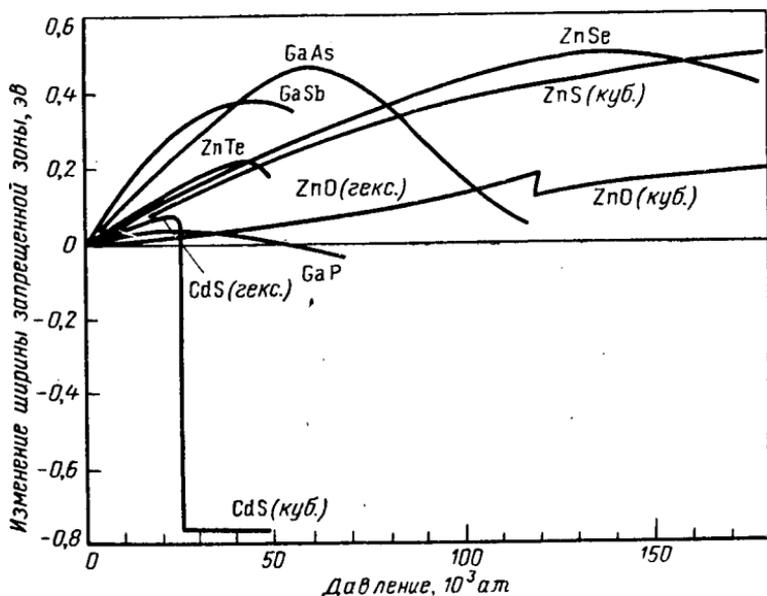


Рис. 10-17. Зависимость ширины запрещенной зоны от давления для некоторых полупроводников.

коротких длин волн. Для кремния $\gamma = 2 \cdot 10^{-6} \text{ эв/кгс/см}^2$, а для германия $\gamma = 8 \cdot 10^{-6} \text{ эв/кгс/см}^2$.

На рис. 10-17 приведены данные, показывающие зависимость ширины запрещенной зоны от давления для некоторых полупроводников.

10-9. СПОНТАННОЕ И ВЫНУЖДЕННОЕ ИЗЛУЧЕНИЕ АТОМОВ

Прежде чем переходить к изучению свечения полупроводников, рассмотрим процесс поглощения и излучения света в атоме. Из всей совокупности квантовых состояний, в которых может находиться

ся атом, для простоты рассмотрения выберем только два состояния (1 и 2) с энергией E_1 и E_2 .

Поглощая квант света $h\nu$, атом из нижнего состояния 1 перейдет в возбужденное 2. Через некоторый промежуток времени dt атом из возбужденного состояния 2 возвратится в нижнее 1, отдавая избыток энергии $E_2 - E_1$ в виде излучения. Переход с излучением может произойти независимо от воздействия внешнего поля, т. е. самопроизвольно, — так называемый спонтанный переход — или под действием поля, в котором находится возбужденный атом. Такой переход называется вынужденным, а излучение носит название вынужденного, индуцированного или стимулированного.

Спонтанные переходы являются случайными, их вероятность обозначим через A_{21} . Если N_2 — число атомов, находящихся в момент времени t в состоянии 2, то число спонтанных переходов $2 \rightarrow 1$ за время dt равно:

$$dN_{21}' = A_{21}N_2 dt. \quad (10-84)$$

Это число переходов равно уменьшению числа атомов, находящихся в момент времени t в возбужденном состоянии, т. е.

$$-dN_2 = A_{21}N_2 dt. \quad (10-85)$$

Интегрируя это уравнение, получим закон убывания числа возбужденных атомов со временем при спонтанных переходах:

$$N_2 = N_{20}e^{-A_{21}t}, \quad (10-86)$$

где N_{20} — число атомов, находящихся в состоянии 2 в момент $t = 0$.

Так как за каждый спонтанный переход излучается энергия $E_2 - E_1 = h\nu$, то за промежуток времени dt выделится энергия, равная

$$h\nu dN_{21}' = A_{21}N_2 h\nu dt. \quad (10-87)$$

Следовательно, интенсивность спонтанного свечения (энергия, испускаемая в единицу времени) равна:

$$I = A_{21}N_{20}e^{-A_{21}t}h\nu = I_0e^{-A_{21}t}, \quad (10-88)$$

где

$$I_0 = A_{21}N_{20}h\nu.$$

Так как $A_{21}N_2dt$ есть число атомов, совершивших за время dt переход $2 \rightarrow 1$, которое определяет количество атомов, «проживших» в возбужденном состоянии время t . Тогда общая продолжительность пребывания в возбужденном состоянии всех атомов, испытавших переход за время от 0 до ∞ , составит:

$$T = \int_0^{\infty} t A_{21}N_2 dt = \int_0^{\infty} N_{20}A_{21} \cdot te^{-A_{21}t} dt, \quad (10-89)$$

откуда следует, что средняя продолжительность жизни атома в возбужденном состоянии равна:

$$\tau = \frac{T}{N_{20}} = \int_0^{\infty} A_{21} t e^{-A_{21}t} dt = \frac{1}{A_{21}}. \quad (10-90)$$

В соответствии с этим формула (10-88) запишется в виде

$$I = I_0 e^{-t/\tau}. \quad (10-91)$$

Таким образом, *интенсивность спонтанного излучения во времени уменьшается по экспоненциальному закону.*

Поскольку спонтанное излучение имеет статистический характер, оно некогерентно. Это обусловлено тем, что акты спонтанного излучения происходят независимо друг от друга в разные моменты времени.

Рассмотрим теперь процессы, идущие в атоме под влиянием внешнего поля. При воздействии излучения, имеющего спектральную плотность энергии u_ν , невозбужденный атом, поглощая энергию поля, перейдет в возбужденное состояние 2. При этом поглощающие переходы совершаются тем чаще, чем больше величина u_ν . И если N_1 — число атомов в невозбужденном состоянии, а $u_\nu B_{12}$ — вероятность перехода с поглощением, то число поглощающих переходов в единицу времени в единице объема будет равно:

$$dN_{12} = B_{12} u_\nu N_1. \quad (10-92)$$

Однако атом, находящийся в возбужденном состоянии (на уровне E_2), под действием поля может отдавать свою энергию, возвращаясь на нижний уровень (на уровень E_1). Этот переход является вынужденным или индуцированным. 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 равно:

$$dN_{21}^* = B_{21} u_\nu N_2. \quad (10-93)$$

Следовательно, *внешнее излучение с частотой ν вызывает не только поглощение света, но и добавочное индуцированное излучение той же частоты.*

Таким образом, переходы с испусканием происходят либо спонтанно, либо индуцированно, и полная вероятность их равна $A_{21} + B_{21} u_\nu$, а среднее число переходов с излучением в единицу времени в единице объема равно:

$$dN_{21} = (A_{21} + B_{21} u_\nu) N_2. \quad (10-94)$$

При термодинамическом равновесии на основании принципа детального равновесия число прямых и обратных процессов в единицу времени одинаково. Следовательно,

$$dN_{12} = dN_{21} \quad (10-95)$$

или

$$B_{12}u_\nu N_1 = (A_{21} + B_{21}u_\nu) N_2. \quad (10-96)$$

Отсюда найдем, что

$$\frac{N_2}{N_1} = \frac{u_\nu B_{12}}{A_{21} + B_{21}u_\nu}. \quad (10-97)$$

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:

$$W_m = C g_m e^{-\frac{E_m}{k_0 T}}, \quad (10-98)$$

где g_m — статистический вес, учитывающий вырождение состояния m .

Поэтому можно записать:

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \frac{e^{-\frac{E_2}{k_0 T}}}{e^{-\frac{E_1}{k_0 T}}}. \quad (10-99)$$

Из сравнения равенств (10-97) и (10-99) следует, что

$$\frac{B_{12}u_\nu}{A_{21} + B_{21}u_\nu} = \frac{g_2}{g_1} e^{-\frac{E_2 - E_1}{k_0 T}}, \quad (10-100)$$

откуда с учетом $E_2 - E_1 = h\nu$ получим:

$$u_\nu = \frac{A_{21}g_2}{\frac{h\nu}{B_{12}g_1 e^{\frac{h\nu}{k_0 T}} - B_{21}g_2}}. \quad (10-101)$$

Для определения коэффициентов A и B используем следующие предельные условия. При $T \rightarrow \infty$ плотность энергии u_ν должна стремиться к бесконечности, поэтому знаменатель в (10-101) стремится к нулю. Из этого следует, что

$$B_{12}g_1 = B_{21}g_2 \quad (10-102)$$

или

$$B_{12} = \frac{g_2}{g_1} B_{21}. \quad (10-103)$$

С учетом этого формула (10-101) запишется в виде

$$u_\nu = \frac{g_2 A_{21}}{g_1 B_{12}} \frac{1}{e^{\frac{h\nu}{k_0 T}} - 1}. \quad (10-104)$$

При $h\nu/k_0T \ll 1$ в первом приближении имеем

$$e^{\frac{h\nu}{k_0T}} = 1 + \frac{h\nu}{k_0T}.$$

Следовательно,

$$u_\nu = \frac{g_2 A_{21}}{g_1 B_{12}} \frac{k_0 T}{h\nu}. \quad (10-105)$$

С другой стороны, при условии $h\nu/k_0T \ll 1$ справедлива формула Рэлея — Джинса¹, по которой

$$u_\nu = \frac{8\pi\nu^2}{c^3} k_0 T. \quad (10-106)$$

Из сопоставления формул (10-105) и (10-106) имеем:

$$\frac{g_2 A_{21}}{g_1 B_{12}} = \frac{8\pi h\nu^3}{c^3}, \quad (10-107)$$

откуда

$$A_{21} = \frac{8\pi h\nu^3}{c^3} \frac{g_1}{g_2} B_{12} = \frac{8\pi h\nu^3}{c^3} B_{21}. \quad (10-108)$$

Индукцированное излучение является когерентным, так как оно происходит в том же направлении, что и излучение, его вызывающее, в одной и той же фазе и с одинаковой поляризацией.

Рассмотрим теперь, при каких условиях может наблюдаться индуцированное свечение. Пусть имеем параллельный пучок монохроматического света интенсивности I . При прохождении слоя вещества толщиной Δx интенсивность света уменьшится на величину $-\Delta I = \alpha \Delta x$. Это уменьшение интенсивности света определяется тем, что в атомах происходят переходы с поглощением энергии в количестве $B_{12} u_\nu h\nu N_1 \Delta x$ и индуцированные переходы с излучением энергии, равной $B_{21} u_\nu h\nu N_2 \Delta x$. В результате можно записать:

$$-\Delta I = \alpha \Delta x = (B_{12} N_1 - B_{21} N_2) u_\nu h\nu \Delta x. \quad (10-109)$$

Из формулы (10-109) 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.

И если учесть, что плотность излучения u_ν в параллельном пучке света сечением 1 см^2 , распространяющегося со скоростью c , равна:

$$u_\nu = \frac{I}{c}, \quad (10-110)$$

то коэффициент поглощения α в соответствии с равенствами (10-109), (10-110) и (10-102) будет иметь вид:

$$\alpha = \frac{h\nu}{c} (B_{12} N_1 - B_{21} N_2) = \frac{h\nu}{c} \left(1 - \frac{g_1 N_2}{g_2 N_1} \right) B_{12} N_1. \quad (10-111)$$

¹ Шпольский Э. В., Атомная физика, т. 1, Гостехиздат, 1949.

Учитывая (10-99), имеем:

$$\alpha = B_{12}N_1 \left(1 - e^{-\frac{E_2 - E_1}{k_0 T}} \right) \frac{h\nu}{c}. \quad (10-112)$$

Так как $E_2 > E_1$, то экспоненциальный член в (10-112) 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 положительное поглощение света, а индуцированные переходы лишь уменьшают его. Причиной этого является то, что в равновесном состоянии число невозбужденных атомов всегда больше числа возбужденных атомов.

Из распределения Больцмана (10-99) следует, что

$$T = - \frac{E_2 - E_1}{k_0 \ln \frac{g_1 N_2}{g_2 N_1}}. \quad (10-113)$$

Поскольку $E_2 > E_1$, а в равновесии $g_1 N_2 < g_2 N_1$, то $T > 0$. Однако если выполняется равенство

$$g_1 N_2 > g_2 N_1 \quad \text{или} \quad \frac{N_2}{N_1} > \frac{g_2}{g_1}, \quad (10-114)$$

то $T < 0$. Таким образом, чтобы создать среду с отрицательным поглощением, необходимо осуществить неравновесное состояние, при котором число возбужденных атомов было бы больше числа атомов, находящихся в нормальном состоянии. Заселенность энергетических уровней атомов, удовлетворяющих неравенству (10-114), носит название инверсной заселенности.

10-10. ЛЮМИНЕСЦЕНЦИЯ ПОЛУПРОВОДНИКОВ

У полупроводников, находящихся в возбужденном состоянии, может наблюдаться испускание электромагнитного излучения — люминесценция. Такое состояние вещества можно осуществить различными способами. При фотолюминесценции излучение появляется за счет поглощаемой энергии света. Свечение, возникающее при бомбардировке образца электронами, называют катодолюминесценцией. Люминесценция может наблюдаться и при возбуждении вещества рентгеновским или γ -излучением, протонами, α - и β -частицами (так называемая радиолюминесценция). Возбуждение может происходить 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 избыток над температурным излу-

чением тела в том случае, 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 разряду неравновесных излучений. Кроме того, это определение выделяет люминесценцию из всех других видов неравновесного излучения, таких как отражение и рассеяние света, тормозное излучение, излучение Вавилова — Черенкова, индуцированное излучение и др.

Люминесценция, как любое электромагнитное излучение, характеризуется интенсивностью, спектральным составом, поляризацией, когерентностью и длительностью послесвечения. И только последний признак отличает люминесценцию от всех других видов неравновесного или вынужденного излучения, которые практически безынерционны; их длительность порядка периода световых колебаний ($\approx 10^{-15}$ сек). В случае люминесценции акты поглощения и излучения света разделены промежуточными процессами, что и приводит к продолжительному существованию свечения после прекращения возбуждения.

У полупроводников различают три вида фотолюминесценции: мономолекулярная, метастабильная и рекомбинационная.

Мономолекулярное и метастабильное излучения твердых тел имеют место в том случае, когда поглощение и излучение света происходят внутри примесного центра.

Рекомбинационное излучение полупроводников возникает при рекомбинации электрона и дырки непосредственно или через центр рекомбинации, который выступает здесь в качестве центра свечения. Рекомбинационное излучение может возникать и при экситонном поглощении света. И в этом случае акты поглощения и излучения разделены процессом передачи энергии в кристалле на расстояния порядка многих постоянных решетки.

10-11. МОНОМОЛЕКУЛЯРНОЕ СВЕЧЕНИЕ ТВЕРДЫХ ТЕЛ

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, во-вторых, их спектр люминесценции, как правило, смещен в длинноволновом направлении по сравнению со спектром поглощения, точнее, сдвинуты максимумы этих полос. 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_b состояний в кристаллической решетке будут квази-молекулярными, как показано на рис. 10-18.

При такой ситуации возможно поглощение света примесным центром, в результате которого электрон с основного уровня U_0 перейдет на возбужденный U_b . 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. 10-18 изображены на оси ординат — шкала 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 величину порядка 10^{-8} сек.

В том случае, когда люминесценция происходит при воздействии на возбужденный примесный центр внешних факторов, сообщающих ему энергию активации, такое свечение называют метастабильным. Это излучение запаздывает относительно возбуждения на время, равное длительности пребывания электрона на метастабильном уровне примеси. Оно непостоянно и зависит от температуры.

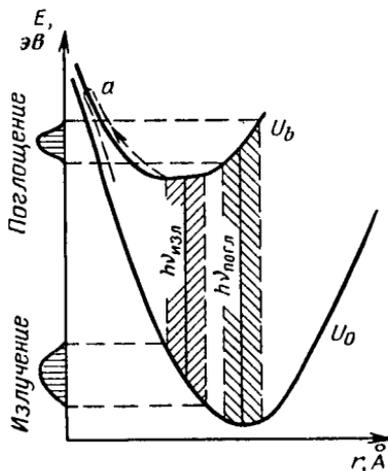


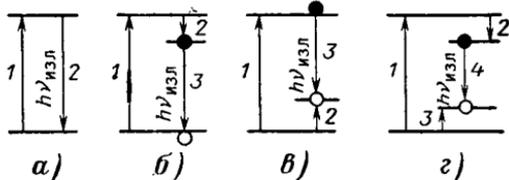
Рис. 10-18. Схема энергетических уровней примесного центра в твердом теле.

10-12. РЕКОМБИНАЦИОННОЕ СВЕЧЕНИЕ ПОЛУПРОВОДНИКОВ

При собственном поглощении света полупроводником имеют место прямые оптические переходы, когда волновой вектор k электрона в валентной зоне совпадает с волновым вектором электрона k' в зоне проводимости ($k' = k$), либо непрямые переходы, когда это условие не выполняется ($k \neq k'$). После процесса поглощения образовавшиеся свободные носители (электрон и дырка) претерпевают рассеяние, которое в течение времени релаксации заставляет электрон опуститься на дно зоны проводимости, а дырку подняться к потолку валентной зоны. Такое равновесное состояние между свободными носителями и кристаллической решеткой устанавливается за время порядка 10^{-10} — 10^{-12} сек. Прямая рекомбинация электрона и дырки с излучением фотона (рис. 10-19, а) наиболее

вероятна, если после процесса релаксации волновые векторы электрона и дырки будут одинаковы ($k = k'$). В частности, такой случай реализуется, когда валентная зона и зона проводимости имеют соответственно максимум и минимум при $k = 0$ (рис. 10-7).

Прямая излучательная рекомбинация электрона и дырки имеет место только в очень чистых и совершенных по кристаллической структуре полупроводниках.



В качестве примера такой люминесценции на рис. 10-20, а представлен спектр рекомбинационного излучения сурьмянистого индия.

Рис. 10-19. Энергетические диаграммы прямой излучательной рекомбинации (а) и излучательной рекомбинации через локальные центры захвата (б, в, г).

В том случае, если у полупроводника имеются прямые и непрямые оптические переходы, в спектре его излучения

наблюдаются два максимума, как, например, в германии (рис. 10-20, б). Длинноволновый максимум излучения определяется непрямыми излучательными переходами с участием фононов, а коротковолновый соответствует прямым переходам.

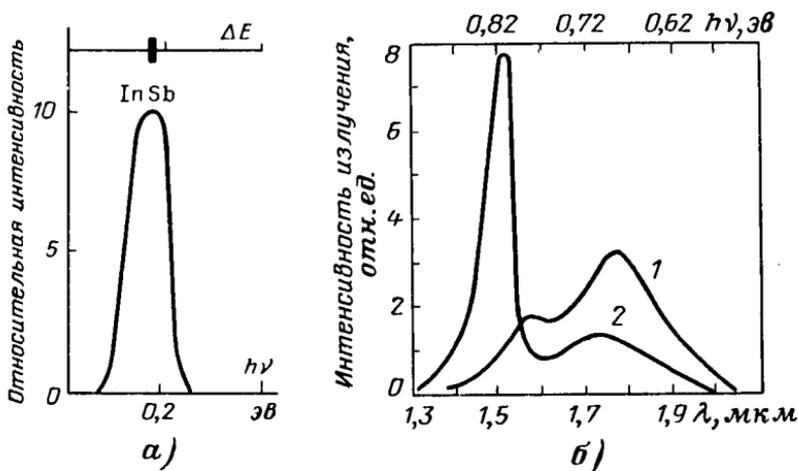


Рис. 10-20. Спектры рекомбинационного излучения антимолида индия (а) и германия (б) при комнатной температуре. У германия экспериментальная кривая (2) дана с поправкой на самопоглощение.

В примесных полупроводниках, в которых в качестве примеси могут выступать не только инородные атомы, но и разного рода дефекты, дислокации, нарушения кристаллической периодичности на поверхности полупроводника, сверхстехиометрический избыток одного из компонентов в случае сложного полупроводникового ма-

териала, важную роль в процессах люминесценции играют локализованные центры, энергетические уровни которых попадают в запрещенную зону полупроводника. В таких случаях люминесценция идет преимущественно через локализованные центры рекомбинации. Уровни электронно-дырочной рекомбинации характеризуются эффективными сечениями захвата электронов σ_n и дырок σ_p . Ловушки и донорные уровни имеют большую вероятность захвата электронов из зоны проводимости, однако обмен с валентной зоной для этих уровней практически невозможен, т. е. $\sigma_n \gg \sigma_p$. Если уровень не занят электроном, то это электронная ловушка, если же уровень занят электроном, то он является донором.

У центра рекомбинации энергетический уровень может быть расположен близко к зоне проводимости (рис. 10-19, б). Такие центры отличаются от ловушек несколько большей глубиной залегания уровней и процессом захвата дырки из валентной зоны. Этот процесс хотя и менее вероятен по сравнению с захватом электрона из зоны проводимости, но все же имеет заметную величину ($\sigma_n > \sigma_p$). Согласно этим предположениям сначала происходит захват электрона из зоны проводимости локальным уровнем (переход 2), а затем осуществляется излучательная рекомбинация локализованного электрона и дырки валентной зоны (переход 3).

В другой модели люминесценции (рис. 10-19, в) локализованный уровень расположен немного выше валентной зоны и имеет большую вероятность захвата дырки ($\sigma_p > \sigma_n$), но вместе с тем имеет и заметную вероятность захвата электронов зоны проводимости. Люминесцентное излучение возникает при рекомбинации свободного электрона зоны проводимости с захваченной этим уровнем дыркой (переход 3).

В третьей модели (рис. 10-19, г) рассматриваются два уровня примеси — основной уровень, расположенный вблизи валентной зоны, и возбужденный уровень, расположенный несколько ниже зоны проводимости. В этом случае основной уровень имеет большую вероятность захвата дырки, а возбужденный — большую вероятность захвата электрона. После того как произошли оба захвата (переходы 2 и 3), происходит переход электрона с возбужденного уровня на основной, сопровождающийся излучением света (переход 4)

10-13. РЕЛАКСАЦИЯ ФОТОЛЮМИНЕСЦЕНЦИИ ПОЛУПРОВОДНИКОВ

При выключении возбуждающего света излучение полупроводника исчезает не мгновенно. Поскольку интенсивность люминесценции пропорциональна числу актов рекомбинации в секунду

$$I_n \sim - \left(\frac{dn}{dt} \right)_r, \quad (10-115)$$

то изменение ее яркости во времени после снятия возбуждения будет различной в зависимости от вида фотолюминесценции.

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

$$n = n_0 e^{-\frac{t}{\tau}}, \quad (10-116)$$

где n_0 — концентрация возбужденных электронов в момент $t = 0$. Следовательно, интенсивность мономолекулярной люминесценции после прекращения возбуждения убывает по экспоненциальному закону (10-91) с постоянной времени τ , являющейся временем жизни электрона на возбужденном уровне центра.

Для рекомбинационной фотолюминесценции закон затухания определяется степенью ее возбуждения. Мы рассмотрим два крайних случая. При слабом возбуждении, когда концентрация неравновесных носителей $\Delta n \ll n_0 + p_0$, интенсивность рекомбинации, согласно выражению (6-18), пропорциональна первой степени концентрации, т. е.

$$\frac{d\Delta n}{dt} = -\frac{\Delta n}{\tau}.$$

Для этого случая $\Delta n = \Delta n_0 e^{-t/\tau}$, а поэтому

$$I_n = I_0 e^{-\frac{t}{\tau}}, \quad (10-117)$$

где Δn_0 и I_0 — концентрация неравновесных носителей и интенсивность люминесценции в момент времени $t_0 = 0$, т. е. в момент выключения возбуждающего света.

Вид излучения полупроводника, интенсивность которого описывается равенством (10-117), носит название мономолекулярного рекомбинационного свечения. По его экспоненциальной кривой релаксации можно определить время жизни неравновесных носителей τ .

В случае сильного возбуждения, когда $\Delta n \gg n_0 + p_0$, интенсивность рекомбинации пропорциональна квадрату концентрации неравновесных носителей [выражение (6-20)], т. е.

$$-\frac{d\Delta n}{dt} = \gamma_r \Delta n^2. \quad (10-118)$$

Поскольку концентрация неравновесных носителей убывает во времени как

$$\Delta n = \frac{\Delta n_0}{1 + \gamma_r \Delta n_0 t}, \quad (10-119)$$

то интенсивность люминесценции описывается соотношением вида

$$I_n = \frac{\gamma_r^2 \Delta n_0^2}{(1 + \gamma_r \Delta n_0 t)^2}, \quad (10-120)$$

а люминесценция называется бимолекулярным рекомбинационным свечением.

10-14. ТЕМПЕРАТУРНОЕ ТУШЕНИЕ ЛЮМИНЕСЦЕНЦИИ ПОЛУПРОВОДНИКОВ

У полупроводников интенсивность люминесценции уменьшается при повышении температуры, т. е. наблюдается температурное тушение излучения. Если через P_r обозначить вероятность излучения света, которая не зависит от температуры, а через P_{nr} — вероятность безызлучательного перехода, то выход люминесценции будет равен отношению

$$\eta = \frac{P_r}{P_r + P_{nr}}. \quad (10-121)$$

Из этого соотношения следует, что температурное тушение люминесценции полупроводников происходит потому, что при увеличении температуры повышается вероятность испускания фононов.

Тушение молекулярного 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_o и U_b (точка a на рис. 10-18). Затем он безызлучательным способом перейдет на основное состояние. Этот переход на рис. 10-18 показан пунктирной стрелкой. Для такого механизма температурного тушения выход люминесценции будет описываться формулой

$$\eta = \frac{1}{1 + C e^{-\frac{E^*}{k_0 T}}}. \quad (10-122)$$

Здесь C — постоянная; E^* — энергия активации, необходимая для перевода электрона в возбужденном состоянии в точку ближайшего расположения потенциальных кривых U_o и U_n .

Температурное тушение люминесценции полупроводника, обусловленной рекомбинацией свободных носителей через локализованный центр, также может описываться соотношением (10-122). Например, центр свечения расположен выше валентной зоны на величину энергии W (рис. 10-19, θ). В этой схеме центр, поглощая квант света, также может возбуждаться. Затем ионизированный 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. В этом случае в формулу (10-122) вместо E^* следует подставить энергию активации W .

В случае излучательной рекомбинации свободных носителей температурное тушение люминесценции можно описать формулой

$$\eta = \frac{1}{1 + CTg}, \quad (10-123)$$

где C и g — константы; T — абсолютная температура.

При выводе этого соотношения использовалась теория прямой безызлучательной рекомбинации электрона зоны проводимости и дырки валентной зоны.

10-15. СТИМУЛИРОВАННОЕ ИЗЛУЧЕНИЕ В ПОЛУПРОВОДНИКАХ

Одним из способов получения среды с отрицательной температурой может явиться создание косвенным путем избытка атомов на более высоком уровне энергии по отношению к более низкому.

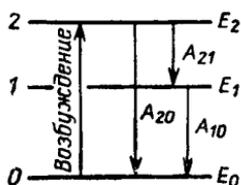


Рис. 10-21. Трехуровневая схема переходов.

Такой случай можно осуществить на атомной или молекулярной системе, которая может находиться в трех энергетических состояниях: в нормальном с энергией E_0 (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} . Будем считать, что эти уровни не имеют вырождения (рис. 10-21).

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 равно $A_{21}N_2$. Разрушается уровень только за счет переходов $1 \rightarrow 0$, число которых равно $A_{10}N_1$.

В стационарных условиях количество актов, заселяющих переходы, равно числу переходов разрушения

$$A_{21}N_2 = A_{10}N_1. \quad (10-124)$$

Для получения индуцированного излучения, как показано в § 10-9, 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

$$A_{21} < A_{10}. \quad (10-125)$$

Из этого соотношения следует, что уровень 2 за счет переходов $2 \rightarrow 1$ должен разрушаться 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 разрушался за счет переходов $2 \rightarrow 0$, т. е. вероятность перехода из состояния 2 в состояние 0 должна быть мала.

Генераторы стимулированного излучения называются 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), содержащую в качестве примеси замещения трехвалентные ионы хрома Cr^{3+} . У трехвалентного иона хрома, энергетическая схема которого изображена на рис. 10-22, 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 $\lambda = 556 \text{ мкм}$. Для состояния 4F_2 среднее время жизни $\sim 5 \cdot 10^{-8} \text{ сек}$, поэтому вероятность перехода ${}^4F_2 \rightarrow {}^2E$ велика, а для перехода ${}^2E \rightarrow {}^4A_2$ мала, так как среднее время жизни для состояния 2E порядка $5 \cdot 10^{-3} \text{ сек}$.

При помощи оптической накачки производится инверсная заселенность метастабильных уровней 2E по отношению к уровню 4A_2 . Переходы ${}^2E \rightarrow {}^4A_2$ сопровождаются двумя линиями излучения с $\lambda_1 = 694,3$ и $\lambda_2 = 692,9 \text{ мкм}$. В спонтанном свечении эти линии слабые, так как вероятность соответствующих переходов мала.

Для наблюдения высвечивания за счет вынужденных переходов кристалл рубина используется в виде стерженька длиной в несколько сантиметров с строго параллельными и оптически отполированными торцами. Торцы стержня покрываются тонким слоем серебра. При этом один из них делается полупрозрачным.

Возникшие линии спонтанного свечения за счет перехода ${}^2E \rightarrow {}^4A_2$ в среде с отрицательным коэффициентом поглощения усиливаются индуцированными переходами. Их интенсивность при прохождении длины стержня \mathcal{L} возрастает по закону

$$I = I_0 e^{\alpha \mathcal{L}}. \quad (10-126)$$

Это усиление будет увеличиваться в результате отражения излучения от торцов кристалла. Многократно отражаться будут лишь те лучи, которые падают на торцы кристалла перпендикулярно. В силу этого излучение, вызванное индуцированными переходами, сосредоточено в пределах малого телесного угла $\Delta\Omega$. В результате возникает узконаправленный поток 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 , а поэтому оно уходит из кристалла через боковые поверхности.

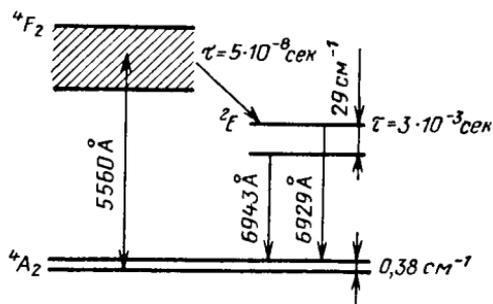


Рис. 10-22. Схема уровней в кристалле рубина.

Индукцированное излучение является когерентным. Оно происходит в том же направлении и имеет ту же фазу, что и первичное. Поэтому лазер можно рассматривать как «резонатор», внутри которого возникает система стоячих волн. Стоячие волны образуются в том случае, если на длине стержня \mathcal{L} уложится целое число полу-волн

$$\frac{\lambda}{2} n = \mathcal{L}, \quad (10-127)$$

где n — целое число.

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 бурно высвечивается в виде вспышки. При работе рубинового лазера в непрерывном режиме требуется 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 оно быстро вводится в систему. В результате возникает лавинообразный процесс индуцированного излучения.

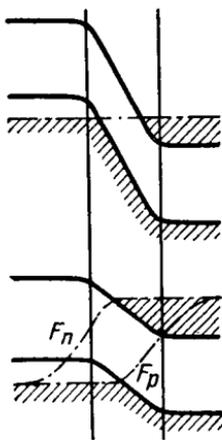


Рис. 10-23. Вырожденный p - n переход в равновесном состоянии и при приложении напряжения в пропускном направлении.

Создать среду с отрицательной температурой также возможно, если искусственно понизить населенность более низкого уровня по отношению к более высокому. Такой механизм возбуждения индуцированного свечения можно осуществить в полупроводниках.

У полупроводников, для которых экстремумы зон находятся в одной и той же точке, например InSb , InAs , InP , GaAs , имеет место преимущественно излучательная рекомбинация. В этом случае при прохождении тока через p - n переход в пропускном направлении рекомбинация инжектированных неосновных носителей сопровождается свечением. Это свечение является спонтанным.

На светящихся 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 должны находиться в вырожденном состоянии. Для вырожденного p - n перехода, энергетическая схема которого изображена на рис. 10-23, достаточно сравнительно небольшого напряжения в пропускном направлении для создания инверсной заселенности.

ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

11-1. ВНУТРЕННИЙ ФОТОЭФФЕКТ

Внутренний фотоэффект — это процесс ионизации атомов полупроводника под действием света, приводящий к образованию добавочных, неравновесных носителей заряда. Добавочную проводимость, обусловленную внутренним фотоэффектом, называют фотопроводимостью.

При внутреннем фотоэффекте первичным процессом является поглощение фотона с энергией, достаточной для возбуждения электрона в зону проводимости (переходы 1 и 2, рис. 11-1) или на локальные уровни, расположенные в запрещенной зоне полупроводника (переход 3). Переход 1 приводит к образованию пары электрон — дырка, тогда как в результате переходов 2 и 3 образуются носители только одного знака.

Если оптическое возбуждение электронов происходит из валентной зоны в зону проводимости, то наблюдается собственная фотопроводимость, которую создают носители обоих знаков. При этом, очевидно, энергия фотона $h\nu$ должна быть не меньше ширины запрещенной зоны полупроводника ($h\nu \geq \Delta E_g = E_c - E_v$).

Для электронов кристаллической решетки справедлив закон сохранения квазиимпульса для прямых и непрямых оптических переходов. При взаимодействии фотона и электрона имеет место прямой (вертикальный) оптический переход (переход 1, рис. 11-2). Однако в кристаллической решетке значительную вероятность имеет и более сложный процесс: взаимодействие фотона, электрона и фонона (кванта колебаний кристаллической решетки). В результате такого взаимодействия электрон приобретает в основном энергию фотона и изменяет свое волновое число за счет фонона (переход 2, рис. 11-2). Такие переходы называют *непрямыми* (невертикальными) оптическими переходами.

обретает в основном энергию фотона и изменяет свое волновое число за счет фонона (переход 2, рис. 11-2). Такие переходы называют *непрямыми* (невертикальными) оптическими переходами.

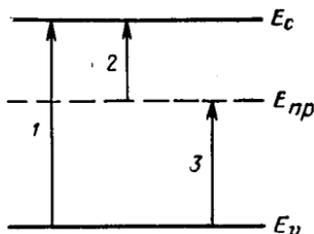


Рис. 11-1. Схема возможных оптических переходов электрона.

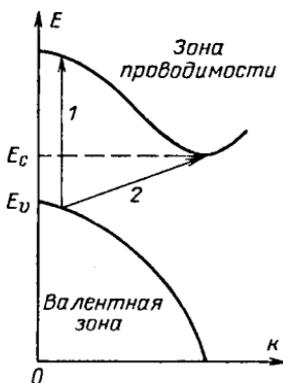


Рис. 11-2. Прямые (1) и непрямые (2) оптические межзональные переходы.

При наличии сложной энергетической зоны прямым оптическим переходам может соответствовать энергия бóльшая, чем энергия термических переходов. Поскольку вероятность непрямых оптических переходов меньше вероятности прямых переходов, то в спектрах поглощения энергии фотонов, соответствующих прямым переходам, должно наблюдаться более или менее резкое возрастание поглощения и, следовательно, фотопроводимости.

Собственная полоса поглощения, всегда имеющая отчетливо выраженную длинноволновую границу, в принципе может иметь и коротковолновую. Однако во многих случаях зона проводимости перекрывается вышележащими разрешенными зонами, образуя сплошной спектр. Поэтому спектр поглощения и спектральная зависимость внутреннего фотозффекта простираются далеко в коротковолновую область. Вместе с тем при больших энергиях фотонов ($h\nu \geq 2\Delta E_g$) фотопереход электрона в зону проводимости может сопровождаться эффектом ударной ионизации, приводящей к освобождению нескольких электронов и дырок. Таким образом, теория внутреннего фотозффекта сводится к теории поглощения лишь в некоторой области спектра вблизи длинноволнового края собственной полосы поглощения.

11-2. ФОТОПРОВОДИМОСТЬ

Неравновесные электроны и дырки, образованные в результате взаимодействия вещества с фотонами достаточно больших энергий, сразу же после процесса ионизации могут иметь энергии значительно бóльшие, чем средняя энергия равновесных носителей, которая равна k_0T . Однако в результате взаимодействия с фононами и дефектами кристаллической решетки неравновесные носители заряда быстро приобретают температуру решетки, и их энергия становится равной средней тепловой энергии равновесных носителей заряда. Этот процесс происходит за время порядка 10^{-10} — 10^{-12} сек, которое равно времени релаксации носителей заряда по энергии. Как правило, время жизни неравновесных носителей заряда значительно превосходит эту величину, составляя 10^{-2} — 10^{-8} сек, и, следовательно, большую часть времени жизни до рекомбинации кинетическая энергия неравновесных носителей заряда соответствует средней тепловой энергии равновесных носителей. Поэтому можно считать, что распределение по энергиям неравновесных носителей заряда в зонах является таким же, как равновесных. Значит, и подвижности неравновесных носителей не отличаются от подвижности равновесных, так как подвижности электронов μ_n и дырок μ_p определяются характером взаимодействия носителей заряда с решеткой и зависят, в частности, от распределения носителей по энергиям.

Итак, генерация носителей заряда под действием света приводит к изменению электропроводности σ полупроводника, которая при наличии неравновесных электронов Δn и дырок Δp может быть

записана в виде

$$\sigma = e[(n_0 + \Delta n) \mu_n + (p_0 + \Delta p) \mu_p], \quad (11-1)$$

где n_0 и p_0 — концентрации равновесных электронов и дырок. Избыточная (неравновесная) проводимость, равная разности проводимостей полупроводника при освещении (σ) и в отсутствие освещения (σ_0), представляет собой фотопроводимость (σ_Φ)

$$\sigma_\Phi = \sigma - \sigma_0 = e(\Delta n \mu_n + \Delta p \mu_p). \quad (11-2)$$

Естественно, что концентрации неравновесных носителей Δn и Δp зависят от интенсивности и длительности освещения полупроводника.

Обозначим количество генерируемых носителей заряда под действием освещения через $\Delta n'$ и $\Delta p'$. Очевидно, что $\Delta n'$ и $\Delta p'$ должны быть пропорциональны световой энергии, поглощаемой в единице объема полупроводника за единицу времени. Если интенсивность монохроматического освещения слоя толщиной dx равна I , а коэффициент поглощения вещества равен α , то количество световой энергии, поглощаемой в единицу времени в единице объема равно:

$$-\frac{dI}{dx} = \alpha I. \quad (11-3)$$

Таким образом, количество генерируемых носителей $\Delta n'$ и $\Delta p'$ пропорционально величине αI . Для области собственного поглощения

$$\Delta n' = \Delta p' = \beta \alpha I. \quad (11-4)$$

Коэффициент пропорциональности β называют коэффициентом квантового выхода, так как он определяет число пар носителей заряда (или число носителей заряда при примесной фотопроводимости), образуемых одним поглощенным квантом света, если интенсивность света I измерять числом квантов в секунду. Обычно квантовый выход β не превышает единицы.

При непрерывном освещении полупроводника светом постоянной интенсивности устанавливается стационарное состояние, характеризующееся постоянной концентрацией неравновесных носителей заряда Δn и Δp . Найдем зависимость Δn и Δp от времени t и определим стационарные значения концентраций неравновесных носителей заряда, считая интенсивность света постоянной во всем объеме образца, что приводит к однородной генерации носителей заряда.

Сразу же после начала освещения, по мере увеличения концентрации неравновесных носителей заряда, начинает нарастать процесс рекомбинации. Поскольку скорость генерации неравновесных носителей заряда остается постоянной при постоянной освещенности, то интенсивность рекомбинации скоро достигает интенсивности процесса генерации носителей и устанавливается стационарное значение неравновесной концентрации фотоносителей $\Delta n_{ст}$ и $\Delta p_{ст}$ (рис. 11-3).

Величины этих концентраций определяются количеством возбужденных светом носителей ($\beta\alpha I$), находящихся в зоне до рекомбинации (в течение времени жизни τ),

$$\Delta n_{ст} = \beta\alpha I\tau_n; \quad (11-5)$$

$$\Delta p_{ст} = \beta\alpha I\tau_p. \quad (11-6)$$

Стационарному значению концентрации неравновесных носителей соответствует стационарное состояние неравновесной проводимости

$$\Delta\sigma_{ст} = \Delta\sigma_{nст} + \Delta\sigma_{pст} = e\beta\alpha I (\tau_n\mu_n + \tau_p\mu_p). \quad (11-7)$$

Здесь α и β характеризуют взаимодействие света с веществом и опи-

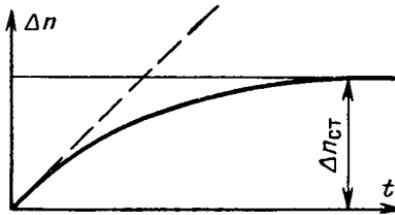


Рис. 11-3. Изменение во времени концентрации носителей заряда при освещении.

сывают процессы генерации неравновесных носителей, а подвижности и времена жизни характеризуют взаимодействие носителей с веществом и определяют процессы движения и рекомбинации неравновесных носителей заряда.

Если один из членов соотношения (11-7) значительно больше другого за счет разницы подвижностей (например, $\mu_n \gg \mu_p$) или времени жизни (например, $\tau_n \gg \tau_p$), то фотопроводимость осуществляется носителем одного знака (в данном случае электронами) и называется **монопольной**

Если один из членов соотношения (11-7) значительно больше другого за счет разницы подвижностей (например, $\mu_n \gg \mu_p$) или времени жизни (например, $\tau_n \gg \tau_p$), то фотопроводимость осуществляется носителем одного знака (в данном случае электронами) и называется **монопольной**

$$\Delta\sigma_{ст} = e\alpha\beta I\tau_n\mu_n. \quad (11-8)$$

11-3. РЕЛАКСАЦИЯ ФОТОПРОВОДИМОСТИ

Пусть полупроводник освещается прямоугольным световым импульсом, как это изображено на рис. 11-4, а.

При включении света стационарное состояние фотопроводимости [выражение (11-7)] достигается не мгновенно, а лишь через некоторое время после начала освещения (рис. 11-4, б). При выключении света неравновесная проводимость исчезнет также через некоторое время после прекращения освещения (рис. 11-4, в). Кривые нарастания и спада неравновесной проводимости называются **кривыми релаксации фотопроводимости**. Нас будут интересовать законы нарастания и спада фотопроводимости в зависимости от уровня освещения. Мы рассмотрим случай малого уровня инжекции и случай большой интенсивности освещения.

Изменение концентрации неравновесных носителей заряда в единицу времени ($d\Delta n/dt$) есть разность между скоростями генерации и рекомбинации носителей

$$\frac{d\Delta n}{dt} = \beta\alpha I - r. \quad (11-9)$$

Второй член правой части уравнения (11-9) учитывает уменьшение концентрации неосновных носителей заряда в результате процесса рекомбинации и согласно (6-15) равен:

$$r = \left(\frac{dn}{dt}\right)_r = \gamma(np - n_0p_0) = \gamma(n_0 + p_0 + \Delta n)\Delta n. \quad (11-10)$$

Интенсивность рекомбинации можно считать пропорциональной концентрации неравновесных носителей только в том случае, если время жизни неравновесных носителей (одинаковое для электронов и дырок) τ не зависит от их концентрации.

Данное условие реализуется, когда концентрация неравновесных носителей Δn , Δp мала по сравнению с концентрацией равновесных основных носителей заряда (например, $\Delta p = \Delta n \ll p_0$), так как при этом изменением концентрации основных носителей под действием освещения можно пренебречь и считать ее постоянной. Этот случай имеет место, например, в примесном полупроводнике при генерации фотоносителей в области фундаментального поглощения при такой температуре, когда вся примесь ионизирована.

Таким образом, при малом уровне освещения, когда $\Delta n \ll \ll n_0 + p_0$, в равенстве (11-10) можно пренебречь величиной Δn , стоящей в скобках. Если обозначить через $1/\tau = \gamma(n_0 + p_0)$, то $r = \Delta n/\tau$. С учетом этого уравнение (11-9) запишется в виде

$$\frac{d\Delta n}{dt} = \beta\alpha I - \frac{\Delta n}{\tau}. \quad (11-11)$$

Найдем решение уравнения (11-11), считая, что $\Delta p = \Delta n \ll p_0$ и $p_0 \gg n_0$ и что полупроводник начинает освещаться в момент времени $t = 0$ светом постоянной интенсивности. Тогда, разделяя переменные и интегрируя с учетом начального условия $\Delta n = 0$ при $t = 0$, получаем:

$$\Delta n = \tau\beta\alpha I \left(1 - e^{-\frac{t}{\tau}}\right). \quad (11-12)$$

Установившееся значение неравновесной концентрации электронов $\Delta n_{ст}$ определяется из уравнения (11-12) при $t \rightarrow \infty$

$$\Delta n_{ст} = \tau\beta\alpha I. \quad (11-13)$$

Если, наоборот, в образце создана стационарная концентрация неравновесных носителей $\Delta n_{ст}$ и в момент $t = 0$ производится

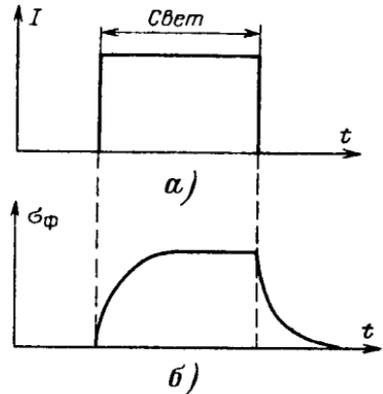


Рис. 11-4. Релаксация неравновесной проводимости (б) при возбуждении прямоугольным импульсом света (а).

выключение света, то концентрация неравновесных носителей заряда спадает до нулевого значения по закону

$$\Delta n = \Delta n_{\text{ст}} e^{-\frac{t}{\tau}} = \tau \beta \alpha I e^{-\frac{t}{\tau}}. \quad (11-14)$$

Таким образом, при малом уровне инжекции релаксация неравновесной концентрации носителей заряда при мгновенном включении и выключении света происходит по экспоненциальному закону с постоянной времени, соответствующей времени жизни неравновесных носителей заряда.

При большом уровне инжекции, когда $\Delta n \gg n_0 + p_0$, для соотношения (11-10) получим:

$$r = -\gamma (\Delta n)^2.$$

В этом случае уравнение (11-9) примет вид:

$$\frac{d\Delta n}{dt} = \beta \alpha I - \gamma (\Delta n)^2. \quad (11-15)$$

Для решения этого уравнения используем начальные условия, аналогичные условиям при малой интенсивности освещения. В результате получим, что для большого уровня инжекции при освещении прямоугольным световым импульсом релаксационные кривые нарастания и спада неравновесной концентрации электронов определяются следующими выражениями:

для нарастания

$$\Delta n = \sqrt{\frac{\beta \alpha I}{\gamma}} \operatorname{th} t \sqrt{\gamma \beta \alpha I}; \quad (11-16)$$

для спада

$$\Delta n = \sqrt{\frac{\beta \alpha I}{\gamma}} \frac{1}{t \sqrt{\gamma \beta \alpha I + 1}}, \quad (11-17)$$

т. е. нарастание описывается гиперболической тангенсоидой, а спад — гиперболой (рис. 11-5).

В этом случае среднее время жизни электронов согласно уравнению (11-10) определяется концентрацией неравновесных дырок в валентной зоне, т. е.

$$\tau = \frac{1}{\gamma \Delta p}. \quad (11-18)$$

Поскольку величина Δp зависит от интенсивности света и времени (в нестационарном случае), то и τ также является переменной величиной. При этом необходимо пользоваться мгновенным значением времени жизни $\tau_{\text{мгн}}$. Для кривой нарастания согласно форму-

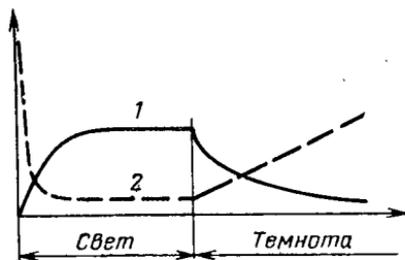


Рис. 11-5. Релаксация неравновесной концентрации электронов (1) и их мгновенное время жизни (2) для случая квадратичной рекомбинации.

лам (11-13) и (11-16) время жизни определяется зависимостью вида

$$\tau_{\text{мгн}} = \frac{1}{\gamma \Delta p} = \frac{1}{\gamma \Delta n} = \frac{1}{\sqrt{\gamma \beta \alpha I}} \operatorname{cth}(t \sqrt{\gamma \beta \alpha I}). \quad (11-19)$$

Для спадающей части релаксационной кривой имеем:

$$\tau_{\text{мгн}} = \frac{1}{\gamma \Delta p} = \frac{1}{\gamma \Delta n} = \frac{1}{\sqrt{\gamma \beta \alpha I}} (t \sqrt{\gamma \beta \alpha I} + 1). \quad (11-20)$$

На рис. 11-5 кривой 2 представлено изменение мгновенного времени жизни в процессе релаксации при большом уровне освещения.

Для описания релаксации фотопроводимости в общем случае можно воспользоваться уравнением (11-11)

$$\frac{d\Delta n}{dt} = \beta \alpha I - \frac{\Delta n}{\tau},$$

если под τ понимать мгновенное значение времени жизни.

Следовательно,

$$\tau_{\text{мгн}} = \frac{\Delta n}{\beta \alpha I - \frac{d\Delta n}{dt}}. \quad (11-21)$$

Для определения времени жизни в стационарном состоянии $\tau_{\text{ст}}$ необходимо в уравнении (11-21) использовать $\Delta n_{\text{ст}}$ и положить $d\Delta n/dt = 0$. Таким образом,

$$\tau_{\text{ст}} = \frac{\Delta n_{\text{ст}}}{\beta \alpha I} \quad (11-22)$$

есть величина постоянная для данной интенсивности света.

Полученные аналитические зависимости для нарастания концентрации носителей заряда позволяют определить закон изменения неравновесной стационарной проводимости (концентрации) от интенсивности освещения, т. е. так называемые люкс-амперные характеристики

$$\Delta n_{\text{ст}} = \tau_{\text{ст}} \beta \alpha I. \quad (11-23)$$

При линейном законе рекомбинации, когда время жизни неравновесных носителей заряда не зависит от интенсивности освещения, люкс-амперная характеристика линейна, так как в соответствии с формулой (11-23) стационарная неравновесная концентрация $\Delta n_{\text{ст}}$ пропорциональна интенсивности света I (кривая 1, рис. 11-6).

При квадратичном законе рекомбинации, когда время жизни неравновесных носителей заряда зависит от интенсивности света,

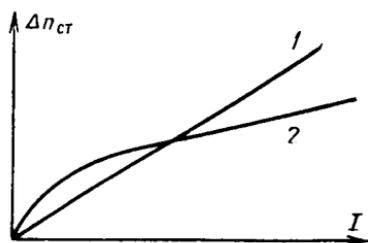


Рис. 11-6. Зависимость стационарной неравновесной концентрации электронов от интенсивности освещения для случая линейной (1) и квадратичной (2) рекомбинации.

на основании (11-19)

$$\tau_{ст} = \frac{1}{\sqrt{\gamma\beta\alpha I}}, \quad (11-24)$$

люкс-амперная характеристика определяется выражением

$$\Delta n_{ст} = \tau_{ст} \beta \alpha I = \sqrt{\frac{\beta \alpha}{\gamma}} I. \quad (11-25)$$

Зависимость $\Delta n_{ст}$ от интенсивности света для случая квадратичной рекомбинации представлена на рис. 11-6 кривой 2.



Рис. 11-7. Кривые релаксации фотопроводимости при отсутствии (1) и при наличии (2) уровней прилипания.

Если в полупроводнике имеются центры прилипания, то они будут оказывать влияние на кинетику фотопроводимости. Электроны зоны проводимости не только будут рекомбинировать с дырками валентной зоны, но и захватываться уровнями прилипания. В результате этого будет иметь место замедление процессов нарастания и спада неравновесной проводимости (рис. 11-7).

11-4. ФОТОПРОВОДИМОСТЬ ПРИ НАЛИЧИИ ПОВЕРХНОСТНОЙ РЕКОМБИНАЦИИ И ДИФФУЗИИ НОСИТЕЛЕЙ ЗАРЯДА

В предыдущем анализе фотопроводимости, где предполагалась однородная генерация носителей заряда по всему объему образца, не была учтена поверхностная рекомбинация носителей заряда, которая приводит к относительному уменьшению концентрации неравновесных носителей заряда вблизи поверхности. Если, как и прежде, предполагать поглощение излучения равномерным, то единственным изменением, связанным с учетом поверхностной рекомбинации, в решении уравнения (11-12) будет замена времени жизни неравновесных носителей заряда τ эффективным временем жизни τ_f (9-52)

$$\frac{1}{\tau_f} = \frac{1}{\tau} + \frac{2s}{d}, \quad (11-26)$$

где s — скорость поверхностной рекомбинации (в предположении, что s невелико); d — размер образца в направлении освещения.

Отношение τ к эффективному времени жизни τ_f характеризует фотопроводимость образца при наличии поверхностной рекомбина-

ции носителей заряда

$$\frac{\tau}{\tau_f} = 1 + \frac{2s\tau}{d}. \quad (11-27)$$

Из формулы (11-27) видно, что если $s\tau \ll d$, то поверхностная рекомбинация слабо влияет на фотопроводимость; если же $s\tau \gg d$, то величина фотопроводимости оказывается в $2s\tau/d$ раз меньше, чем при $s = 0$, и определяется только условиями рекомбинации носителей заряда на поверхности

$$\frac{1}{\tau_f} \approx \frac{2s}{d}. \quad (11-28)$$

Наличие процесса поверхностной рекомбинации может оказать влияние не только на стационарное значение фотопроводимости,

но и на ее спектральную зависимость. Качественно это можно пояснить следующим образом.

При изменении спектральной зависимости фотопроводимости в области края фундаментального поглощения, где коэффициент поглощения $\alpha(\lambda)$ сильно возрастает (до $10^4 - 10^5 \text{ см}^{-1}$), необходимо учитывать два конкурирующих процесса. Один из них приводит к увеличению фотопроводимости за счет роста коэффициента поглощения $\alpha(\lambda)$ (при условии, что образец освещается монохроматическим светом при постоянном потоке фотонов). Второй процесс состоит в относительном увеличении вблизи поверхности концентрации электронов, которые имеют меньшее время жизни (11-28), чем носители в объеме. Эти процессы могут привести к тому, что в спектральной зависимости фотопроводимости будет наблюдаться максимум на краю фундаментального поглощения (рис. 11-8). Очевидно, чем больше скорость поверхностной рекомбинации s , тем сильнее будет выражен максимум фотопроводимости.

Теперь рассмотрим случай, когда падающее излучение поглощается неравномерно по толщине d образца и имеет место диффузия носителей заряда. Будем считать образец толстым, так что $d \gg 1/\alpha$ и $d \gg L$ (L — биполярная диффузионная длина). Условие $d \gg 1/\alpha$ указывает на то, что почти все излучение поглощается в тонком поверхностном слое образца и интенсивностью излучения, отраженного от неосвещенной поверхности образца, можно пренебречь. В этом случае интенсивность излучения I' внутри образца описывается уравнением

$$I' = I(1 - R_s)e^{-\alpha x}, \quad (11-29)$$

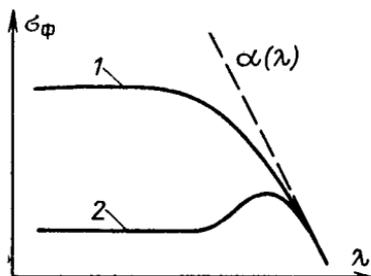


Рис. 11-8. Спектральная зависимость фотопроводимости (2) вблизи края фундаментального поглощения (1).

где R_s — коэффициент отражения от поверхности полубесконечного образца; I — интенсивность падающего света.

Уравнение для стационарной концентрации избыточных носителей заряда $\Delta p(x)$ в образце электронной проводимости при наличии неравномерной генерации и диффузии носителей заряда имеет вид:

$$D \frac{d^2 \Delta p}{dx^2} - \frac{\Delta p}{\tau} + g e^{-\alpha x} = 0, \quad (11-30)$$

где D — коэффициент биполярной диффузии электронов и дырок;

$$g = \beta \alpha I (1 - R_s).$$

Соответствующим решением уравнения (11-30) будет функция

$$\Delta p(x) = A e^{-\frac{x}{L}} + B e^{-\alpha x}, \quad (11-31)$$

где $L = \sqrt{D\tau}$ — биполярная диффузионная длина.

Постоянный коэффициент B находим при подстановке выражения (11-31) в уравнение (11-30)

$$B = -\frac{g\tau}{L^2\alpha^2 - 1}. \quad (11-32)$$

Коэффициент A можно определить из граничного условия на освещенной поверхности (при $x = 0$)

$$D \frac{d\Delta p}{dx} = s\Delta p, \quad (11-33)$$

означающего, что диффузионный поток избыточных носителей к освещенной поверхности пропорционален концентрации неравновесных носителей.

Поскольку роль поверхностной рекомбинации в явлении фотопроводимости была оценена ранее, положим в данном случае скорость поверхностной рекомбинации $s = 0$ и определим константу A из условия

$$\frac{d\Delta p}{dx} = 0. \quad (11-34)$$

Тогда распределение носителей заряда по образцу можно записать в виде

$$\Delta p(x) = \frac{g\tau}{\alpha^2 L^2 - 1} \left(\alpha L e^{-\frac{x}{L}} - e^{-\alpha x} \right). \quad (11-35)$$

Необходимо отметить, что если $\alpha L \gg 1$, то избыточные носители проникают на значительно большую глубину, чем излучение, и глубина их проникновения по порядку величины равна диффузионной длине L . Интегрируя уравнение (11-35) по x , можно определить полную концентрацию избыточных носителей заряда Δp , принимающих участие в фотопроводимости образца в направлении

y (третий размер образца в направлении z считаем равным 1):

$$\begin{aligned} \Delta p &= \int_0^d \Delta p(x) dx = \int_0^{\infty} \Delta p(x) dx = \\ &= \frac{\tau g}{\alpha^2 L^2 - 1} \int_0^{\infty} (\alpha L e^{-\frac{x}{L}} - e^{-\alpha x}) dx = \tau \beta I (1 - R_s). \end{aligned} \quad (11-36)$$

Из выражения (11-36) видно, что концентрация избыточных носителей пропорциональна полному числу фотонов, поглощенных единицей поверхности образца.

11-5. ПРИМЕСНАЯ ФОТОПРОВОДИМОСТЬ

При наличии в запрещенной зоне полупроводника локальных примесных уровней излучение, так же как и тепловое возбуждение, может вызывать переходы электронов между примесными уровнями и зонами (рис. 11-9, *а*, переходы 2, 3). Поглощение и фотопроводи-

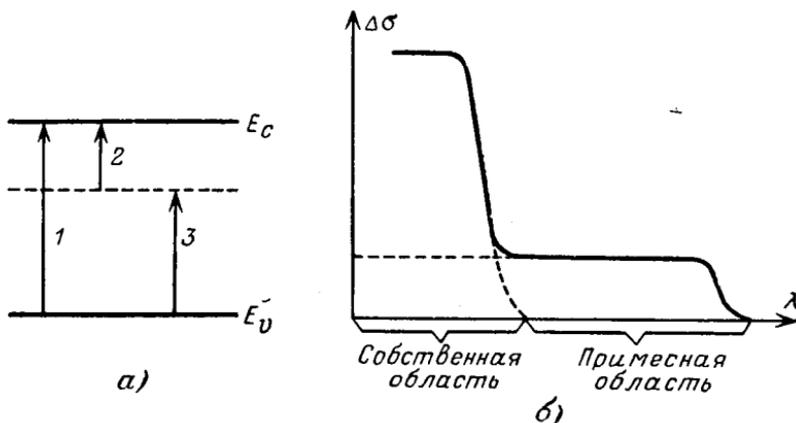


Рис. 11-9. Схематическое изображение переходов (*а*) и спектрального распределения фотопроводимости (*б*).

мость, обусловленные такими переходами, называют примесными. Энергия ионизации примеси, уровни которой расположены в запрещенной зоне, естественно, меньше, чем ширина запрещенной зоны ΔE_g . Поэтому длинноволновая граница примесного поглощения и фотопроводимости сдвинута в длинноволновую сторону спектра по отношению к собственному поглощению и фотопроводимости. Кроме того, интенсивность примесной фотопроводимости значительно меньше собственной, так как концентрация примеси на много порядков меньше концентрации атомов основной решетки. Схематическое изображение спектрального распределения фотопроводимости представлено на рис. 11-9, *б*.

На рис. 11-10 показаны спектральные зависимости фотопроводимости германия с различными примесями. Энергия, соответствующая длинноволновой границе фотопроводимости примеси, близка к энергии ее термической ионизации.

В случае примесного поглощения интенсивность генерации носителей $\beta \alpha I$ изменяется нелинейно с изменением интенсивности света, так как коэффициент поглощения света α в примесной области поглощения не остается постоянным, а уменьшается с увеличением

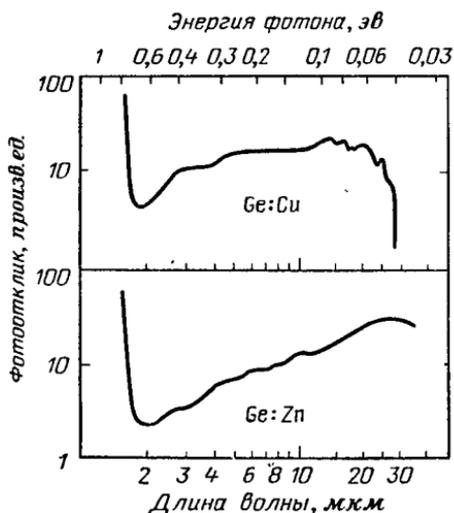


Рис. 11-10. Спектральные кривые фотопроводимости для кристаллов германия, легированных медью и цинком.

интенсивности света из-за заметного опустошения примесных центров. Поэтому люкс-амперная характеристика в области примесной фотопроводимости будет линейна лишь при малых интенсивностях света и будет достигать насыщения для больших интенсивностей света при полном опустошении примесных центров.

Зависимость коэффициента поглощения α (I) в примесной области от интенсивности излучения имеет также важное значение в процессах релаксации примесной фотопроводимости. В простом случае возбуждение примесной фотопроводимости связано с появлением свободных носителей только одного знака, тогда как заряды противоположного знака остаются локализованными на атомах примеси.

Поэтому процессы совместной диффузии и дрейфа пар электрон — дырка в условиях электронейтральности не могут иметь места. Характеристикой процесса диффузии является не диффузионная длина, а дебаевская длина экранирования. При концентрациях носителей заряда $n_0 \approx 10^{17} - 10^{18} \text{ см}^{-3}$, характерных для таких материалов, как германий и кремний, дебаевская длина экранирования весьма мала ($10^{-8} - 10^{-4} \text{ см}$) и, следовательно, процессы диффузии носителей заряда при их неоднородной генерации можно не учитывать.

11-6. ЭФФЕКТ ДЕМБЕРА

Если однородный полупроводник осветить сильно поглощаемым светом, то в его приповерхностном слое, где происходит поглощение света, возникнут неравновесные электроны и дырки, которые будут диффундировать в глубь полупроводника (рис. 11-11). Электроны

опережают дырки, так как их подвижность больше, чем дырок. Это приведет к разделению зарядов, и верхняя грань приобретает положительный заряд, а нижняя — отрицательный, т. е. в полупроводнике возникнет электрическое поле и э. д. с. Эту объемную э. д. с., возникающую при неравномерном освещении однородного полупроводника, называют э. д. с. Дембера.

Проведем вычисление стационарного поля и э. д. с. Дембера для изолированного полупроводника с биполярной проводимостью. Рассмотрение задачи проведем для одномерного случая. В этом случае электронная и дырочная составляющие тока запишутся в виде

$$j_n = en\mu_n \mathcal{E}_{ст} + \mu_n k_0 T \frac{dn}{dx}; \quad (11-37)$$

$$j_p = e p \mu_p \mathcal{E}_{ст} - \mu_p k_0 T \frac{dp}{dx}, \quad (11-38)$$

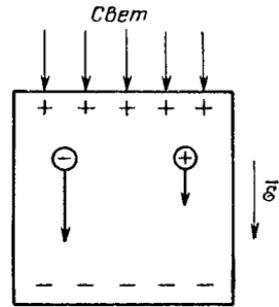


Рис. 11-11. Возникновение э. д. с. Дембера.

где $\mathcal{E}_{ст}$ — внутреннее электрическое поле, возникающее из-за отклонения от нейтральности. Суммарный ток

$$j = j_n + j_p = e(n\mu_n + p\mu_p) \mathcal{E}_{ст} + k_0 T \left(\mu_n \frac{dn}{dx} - \mu_p \frac{dp}{dx} \right). \quad (11-39)$$

В случае изолированного полупроводника $j = 0$, и для поля Дембера получим следующее выражение:

$$\mathcal{E}_{ст} = - \frac{k_0 T}{e} \frac{(\mu_n \frac{dn}{dx}) - \mu_p \frac{dp}{dx}}{n\mu_n + p\mu_p}. \quad (11-40)$$

В случае монополярной проводимости, например $p = 0$, имеем:

$$\mathcal{E}_{ст} = - \frac{k_0 T}{e} \frac{1}{n} \frac{dn}{dx}, \quad (11-41)$$

и э. д. с. Дембера между точками x_1 и x_2 , для которых концентрация электронов n_1 и n_2 , равна:

$$|V_{12}| = \frac{k_0 T}{e} \int_{x_1}^{x_2} \frac{dn}{n} = \frac{k_0 T}{e} \ln \frac{n_2}{n_1}. \quad (11-42)$$

Расчет э. д. с. Дембера для биполярного случая проведем в предположении, что в полупроводнике мало ловушек и нет поверхностной рекомбинации. При этих условиях можно считать, что избыточные концентрации электронов и дырок равны, т. е.

$$\Delta n = \Delta p. \quad (11-43)$$

Поскольку $\frac{dn}{dx} = \frac{dp}{dx} = \frac{d\Delta n}{dx}$,

то

$$|V_{12}| = \frac{k_0 T}{e} \int_{x_1}^{x_2} \frac{\mu_n - \mu_p}{n\mu_n + p\mu_p} d\Delta n. \quad (11-44)$$

Так как $\sigma = e(n\mu_n + p\mu_p)$, а $\Delta\sigma = e(\mu_n + \mu_p)\Delta n$, то

$$\frac{d\Delta n}{dx} = \frac{1}{e(\mu_n + \mu_p)} \frac{d\Delta\sigma}{dx}.$$

Следовательно,

$$\mathcal{E}_{\text{ср}} = \frac{k_0 T}{e} \frac{\mu_p - \mu_n}{\mu_n + \mu_p} \frac{1}{\sigma} \frac{d\Delta\sigma}{dx}, \quad (11-45)$$

а э. д. с. Дембера равна:

$$|V_{12}| = \frac{k_0 T}{e} \frac{\mu_n - \mu_p}{\mu_n + \mu_p} \int_{x_1}^{x_2} \frac{d\Delta\sigma}{\sigma} = \frac{k_0 T}{e} \frac{\mu_n - \mu_p}{\mu_n + \mu_p} \ln \frac{\sigma_2}{\sigma_1}. \quad (11-46)$$

Из выражения (11-46) следует, что э. д. с. Дембера тем больше, чем сильнее различаются подвижности электронов и дырок. При $\mu_n = \mu_p$ э. д. с. Дембера равна нулю.

При низком уровне освещения σ_2 мало отличается от σ_1 , т. е. $\frac{\sigma_2 - \sigma_1}{\sigma} \ll 1$, где $\sigma = \frac{\sigma_1 + \sigma_2}{2}$. Пусть $\delta\sigma = \sigma_2 - \sigma_1$, разлагая выражение $\ln \frac{\sigma_2}{\sigma_1}$ в ряд и ограничиваясь первым членом разложения, получаем:

$$|V_{12}| = \frac{k_0 T}{e} \frac{\mu_n - \mu_p}{\mu_n + \mu_p} \frac{\delta\sigma}{\sigma}. \quad (11-47)$$

Таким образом, при малом уровне возбуждения э. д. с. Дембера между двумя точками образца прямо пропорциональна разности проводимости в этих точках и тем больше, чем меньше среднее значение проводимости полупроводника.

Подсчитаем величину э. д. с. Дембера для Ge с удельной проводимостью $\sigma = 1 \text{ ом}^{-1} \cdot \text{см}^{-1}$ при $T = 300^\circ \text{ К}$, для которого $\mu_n = 3900 \text{ см}^2/\text{сек}$ и $\mu_p = 1900 \text{ см}^2/\text{сек}$. При $\delta\sigma = 0,2 \text{ ом}^{-1} \cdot \text{см}^{-1}$ получаем:

$$V_{12} = 0,025 \frac{2000 \cdot 0,2}{5800} \approx 1,7 \cdot 10^{-3} \text{ в} = 1,7 \text{ мв.}$$

11-7. ФОТОМАГНИТНОЭЛЕКТРИЧЕСКИЙ ЭФФЕКТ

Если неравномерно освещенный полупроводник поместить в магнитное поле, перпендикулярное направлению распространения света, то в нем возникает э. д. с. (рис. 11-12). Это явление называется фотомангнитноэлектрическим эффектом.

На диффундирующие электроны и дырки, возбужденные светом, в магнитном поле действует сила Лоренца, отклоняющая их в

направлении, перпендикулярном их движению. При этом потоки электронов и дырок отклоняются в разные стороны. Вследствие этого у противоположных сторон образца, как это представлено на рис. 11-13, накапливаются заряды противоположного знака, т. е. в направлении, перпендикулярном распространению света и магнитного поля, возникают электрическое поле и э. д. с. Накопление зарядов и возрастание э. д. с. будет продолжаться до тех пор,

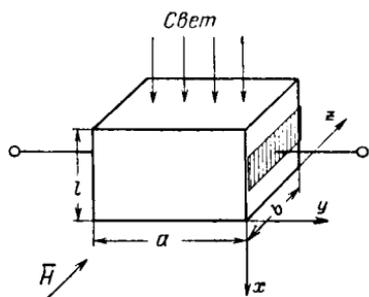


Рис. 11-12. Возникновение фотомагнитной э. д. с.

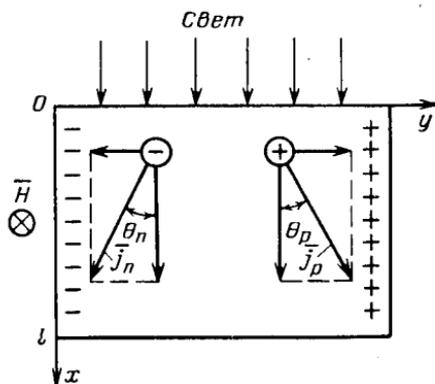


Рис. 11-13. Действие магнитного поля на диффузионные токи электронов и дырок.

пока ток проводимости, обусловленный возникшим электрическим полем, не скомпенсирует магнитно-диффузионный ток.

Определение фотомагнитной э. д. с. и тока короткого замыкания проведем для полупроводника в случае биполярной проводимости. Кроме того, будем считать, что скорость поверхностной рекомбинации мала, а поглощение света происходит в приповерхностном слое образца.

В случае слабого магнитного поля ($\mu B \ll 1$) углы Холла для электронов φ_n и для дырок φ_p согласно выражениям (1-34) и (1-35) равны:

$$\varphi_n = -\mu_n^{(n)} B, \quad (11-48)$$

$$\varphi_p = \mu_p^{(n)} B, \quad (11-49)$$

где $\mu_n^{(n)}$ и $\mu_p^{(n)}$ — холловские подвижности электронов и дырок. Они связаны с дрейфовыми подвижностями соотношениями

$$\mu_n^{(n)} = r\mu_n \quad \text{и} \quad \mu_p^{(n)} = r\mu_p,$$

где r — коэффициент, зависящий от механизма рассеяния (§ 5-3).

Если j_{nx} и j_{px} есть диффузионные токи в направлении оси x , то магнитно-диффузионные токи согласно рис. 11-13 равны:

$$j_{ny} = j_{nx} \operatorname{tg} \varphi_n = -r\mu_n B j_{nx}; \quad (11-50)$$

$$j_{py} = j_{px} \operatorname{tg} \varphi_p = r\mu_p B j_{px}. \quad (11-51)$$

Полный магнитно-диффузионный ток равен

$$j_y = j_{ny} + j_{py} = rB (\mu_p j_{px} - \mu_n j_{nx}). \quad (11-52)$$

Если в полупроводнике нет ловушек, то $\Delta n = \Delta p$. С учетом этого диффузионные токи запишутся в виде

$$j_{px} = -j_{nx} = -eD \frac{d\Delta n}{dx}, \quad (11-53)$$

где $D = \frac{n+p}{\frac{n}{D_p} - \frac{p}{D_n}}$ — коэффициент биполярной диффузии.

Подставляя выражение (11-53) в (11-52), получаем:

$$j_y = -erDB (\mu_n + \mu_p) \frac{d\Delta n}{dx}. \quad (11-54)$$

Для того чтобы найти ток короткого замыкания во внешней цепи j_{ks} , необходимо подсчитать суммарный магнитно-диффузионный ток, текущий через всю боковую поверхность (сечение образца в плоскости xz). Если D считать величиной постоянной, а b — размер образца по оси z , то

$$\begin{aligned} j_{ks} &= \int_0^l j_y b dx = -erDB (\mu_n + \mu_p) b \int_0^l d\Delta n = \\ &= erDBb (\mu_n + \mu_p) [\Delta n(0) - \Delta n(l)]. \end{aligned} \quad (11-55)$$

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 фотомагнитной э. д. с. $V_{\text{фм}}$ и полной проводимостью образца σ :

$$j = V_{\text{фм}} \sigma = j_{ks}. \quad (11-56)$$

Поскольку полная проводимость образца равна:

$$\begin{aligned} \sigma &= \frac{bl}{a} \sigma_0 + \frac{b}{a} \int_0^l \Delta \sigma dx = \\ &= \frac{bl}{a} \sigma_0 + \frac{eb}{a} (\mu_n + \mu_p) \int_0^l \Delta n dx, \end{aligned} \quad (11-57)$$

где a — размер образца по оси y ; σ_0 — проводимость в темноте, то полная величина фотомагнитной э. д. с. равна:

$$V_{\text{фм}} = \frac{j_{ks}}{\sigma} = a \frac{erDB (\mu_n + \mu_p) [\Delta n(0) - \Delta n(l)]}{l\sigma_0 + e (\mu_n + \mu_p) \int_0^l \Delta n dx}. \quad (11-58)$$

Следовательно, фотомагнитная э. д. с. пропорциональна магнитному полю и определяется распределением неравновесных носителей Δn .

Если генерация носителей заряда происходит вблизи освещенной поверхности, то концентрация неравновесных носителей заряда из-за диффузии вдоль оси x меняется по закону

$$\Delta n = \Delta n(0) e^{-\frac{x}{L_D}}. \quad (11-59)$$

В том случае, если толщина образца l значительно превосходит диффузионную длину L_D , т. е. $l \gg L_D$, то $\Delta n(l) = 0$, а

$$\int_0^l \Delta n dx = \Delta n(0) \int_0^l e^{-\frac{x}{L_D}} dx = \Delta n(0) L_D. \quad (11-60)$$

Следовательно,

$$V_{\text{фм}} = \frac{earDB (\mu_n + \mu_p) \Delta n(0)}{l\sigma_0 + e (\mu_n + \mu_p) \Delta n(0) L_D}. \quad (11-61)$$

При высоком уровне возбуждения, когда $\Delta n \gg n_0 + p_0$, в равенстве (11-61) можно пренебречь членом $l\sigma_0$, а поэтому будем иметь:

$$V_{\text{фм}} = arDB \frac{1}{L_D}. \quad (11-62)$$

При низком уровне возбуждения, когда $\Delta n \ll n_0 + p_0$, в знаменателе выражения (11-61) можно пренебречь вторым слагаемым. В результате получим:

$$V_{\text{фм}} = earDB \frac{(\mu_n + \mu_p) \Delta n(0)}{e\sigma_0}. \quad (11-63)$$

В случае, если скорость поверхностной рекомбинации пренебрежимо мала, полное число неравновесных носителей заряда в образце, приходящихся на единицу освещенной поверхности, определяется квантовым выходом β , интенсивностью возбуждающего света I и временем жизни носителей τ . Следовательно, учитывая выражение (11-60), можно написать:

$$\int_0^l \Delta n dx = \beta\tau I = \Delta n(0) L_D, \quad (11-64)$$

откуда

$$\Delta n(0) = \frac{\beta\tau I}{L_D} = \frac{\beta I L_D}{D}. \quad (11-65)$$

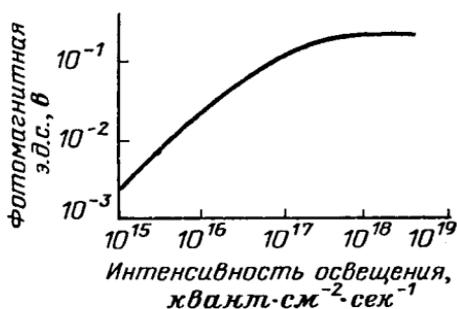


Рис. 11-14. Зависимость фотомангнитной э. д. с. от интенсивности освещения для германия.

Подставляя выражение (11-65) в (11-63), получаем:

$$V_{\text{фм}} = e a r B \frac{(\mu_n + \mu_p) \beta I L_D}{e \sigma_0}. \quad (11-66)$$

Из анализа формул (11-66) и (11-62) следует, что *фотомагнитная э. д. с. при малом уровне возбуждения пропорциональна интенсивности света, а с ростом уровня освещения она стремится к насыщению* (рис. 11-14).

11-8. ВЕНТИЛЬНАЯ ФОТОЭЛЕКТРОДВИЖУЩАЯ СИЛА

При освещении *p-n* перехода светом возникает фото-э. д. с., которая носит название *вентильной э. д. с.* На *p-n* переходе, как известно (§ 8-8), возникает контактная разность потенциалов.

Этот потенциальный барьер обусловлен электрическим полем, которое появляется в результате диффузии носителей заряда через *p-n* переход. При термодинамическом равновесии положение уровня Ферми во всей системе постоянно; энергетическая схема *p-n* перехода изображена на рис. 11-15, а.

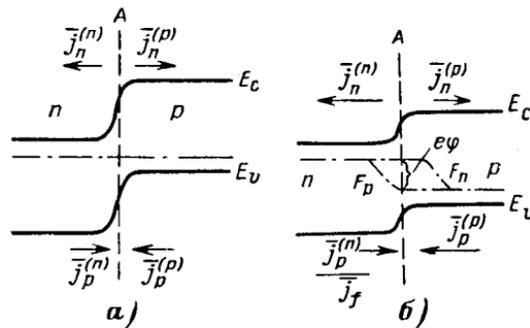


Рис. 11-15. Энергетическая схема *p-n* перехода и токи носителей заряда при термодинамическом равновесии (а) и при освещении (б).

При рассмотрении явлений на *p-n* переходе будем исходить из диодной теории (§ 8-9). В этом

случае токи обусловлены тепловыми процессами, и в равновесии сумма тепловых токов равна нулю.

Через *p-n* переход протекают токи электронов $j_n^{(n)}$ и дырок $j_p^{(n)}$ из *n*-области в *p*-область, а из *p*-области в *n*-область — токи $j_p^{(p)}$ и $j_n^{(p)}$. В принятом обозначении направления токов, как это сделано на рис. 11-15, условие равновесия запишется следующим образом:

$$-j_n^{(n)} + j_n^{(p)} - j_p^{(p)} + j_p^{(n)} = 0. \quad (11-67)$$

В этом равенстве каждая пара токов электронов и дырок равна нулю

$$-j_n^{(n)} + j_n^{(p)} = 0 \quad \text{и} \quad -j_p^{(p)} + j_p^{(n)} = 0, \quad (11-68)$$

так как количества переходящих носителей заряда в прямом и обратном направлении при термодинамическом равновесии равны. Но потоки неосновных носителей — 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. Полный ток насыщения равен:

$$j_s = j_{ns} + j_{ps} = \frac{eD_n n_p}{L_n} + \frac{eD_p p_n}{L_p}. \quad (11-69)$$

Рассмотрим теперь поведение p - n перехода при освещении одной из областей, например электронной, как это изображено на рис. 11-15, б.

Под воздействием света в электронном полупроводнике создаются пары электрон — дырка. В том случае, когда расстояние от места возникновения пар до p - n перехода составляет величину, меньшую длины диффузионного смещения, пары вследствие диффузии подходят к переходу, где они разделяются под воздействием контактного поля. Дырки, для которых нет потенциального барьера, переводятся контактным полем в дырочный полупроводник. Добавочный дырочный ток, обусловленный освещением, обозначим через j_f . Изменение же концентрации электронов можно не учитывать, ибо неравновесные электроны быстро принимают распределение, соответствующее температуре образца, и лишь очень малая их доля, преодолевая потенциальный барьер, перейдет в p -полупроводник. Добавочный дырочный ток, идущий из электронного полупроводника, заряжает дырочный полупроводник положительно относительно электронного, и, следовательно, энергетические уровни электронов в p -области снижаются (рис. 11-15, б).

Таким образом, разделение зарядов приводит к появлению дополнительной составляющей электрического поля, а следовательно, к возникновению на переходе разности потенциалов, представляющей собой для внешней цепи фото-э. д. с. Вместе с тем происходит изменение и потенциального барьера, существующего в неосвещенном p - n переходе. Возникшая фото-э. д. с. уменьшает этот барьер, что в свою очередь приводит к появлению дополнительных потоков основных носителей, направленных навстречу друг другу. При этом токи направлены в одну сторону: из p - в n -область (рис. 11-15, б). Эти токи по своему характеру подобны протекающему через переход току, который имеет место при приложении к переходу пропускного напряжения.

Стационарное состояние установится тогда, когда число создаваемых светом пар электрон — дырка сравняется с числом носителей заряда, уходящих за счет потока через p - n переход и во внешнюю нагрузку, по которой в результате этого потечет ток j . По отношению ко внешней цепи освещенный полупроводник, имеющий p - n переход, будет являться фотозлементом.

Для определения вентильной фото-э. д. с., соответствующей стационарному состоянию, можно написать:

$$j_f - j_n^{(n)} + j_n^{(p)} - j_p^{(p)} + j_p^{(n)} = j. \quad (11-70)$$

Здесь токи неосновных носителей при освещении $j_n^{(p)}$ и $j_p^{(n)}$, как следует из рис. 11-15, б, равны своим значениям в равновесии

$$j_n^{(p)} = j_{ns} \quad \text{и} \quad j_p^{(n)} = j_{ps}. \quad (11-71)$$

Токи основных носителей при освещении $j_n^{(n)}$ и $j_p^{(p)}$ в результате уменьшения потенциала барьеров на $e\varphi$ увеличиваются и становятся равными

$$j_n^{(n)} = j_{ns} e^{\frac{e\varphi}{k_0 T}}; \quad (11-72)$$

$$j_p^{(p)} = j_{ps} e^{\frac{e\varphi}{k_0 T}}. \quad (11-73)$$

Подставив уравнения (11-71) — (11-73) в выражение (11-70), получим:

$$j_f - j_{ns} \left(e^{\frac{e\varphi}{k_0 T}} - 1 \right) - j_{ps} \left(e^{\frac{e\varphi}{k_0 T}} - 1 \right) = j \quad (11-74)$$

или

$$j_f = j_s \left(e^{\frac{e\varphi}{k_0 T}} - 1 \right) + j,$$

откуда для разности потенциалов на p - n переходе будем иметь:

$$\varphi = \frac{k_0 T}{e} \ln \left[1 + \frac{j_f - j}{j_s} \right]. \quad (11-75)$$

Уравнение (11-75) является общим уравнением фотодиода, пригодным для любого режима.

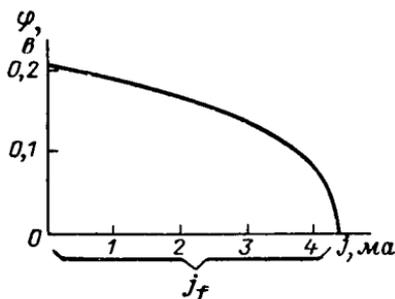


Рис. 11-16. Вольт-амперная характеристика германиевого фотодиода в вентильном режиме.

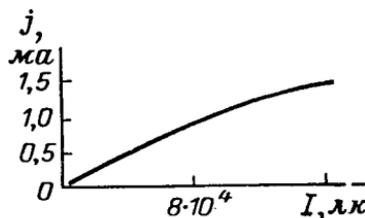


Рис. 11-17. Люкс-амперная характеристика германиевого фотодиода.

Для определения вентильной фото-э. д. с., которая соответствует напряжению на зажимах разомкнутой цепи, необходимо положить $j = 0$. Тогда

$$\varphi_{\text{вен}} = \frac{k_0 T}{e} \ln \left(1 + \frac{j_f}{j_s} \right). \quad (11-76)$$

Величина j_f определяется числом неравновесных дырок, созданных светом и дошедших до p - n перехода. Если через I обозначим число квантов света, падающих каждую секунду на единицу поверхности, β — квантовый выход их, т. е. количество пар носи-

телей, возникающих на один падающий квант света, а через γ — долю непрорекомбинировавших пар носителей заряда, пришедших к p - n переходу, то

$$j_f = e\gamma\beta I. \quad (11-77)$$

Следовательно,

$$\Phi_{\text{вент}} = \frac{k_0 T}{e} \ln \left(1 + \frac{e\gamma\beta I}{j_s} \right). \quad (11-78)$$

При большом уровне освещения, когда $j_f/j_s \gg 1$, имеем:

$$\Phi_{\text{вент}} = \frac{k_0 T}{e} \ln \frac{e\gamma\beta I}{j_s}. \quad (11-79)$$

При малом уровне возбуждения, когда $j_f/j_s \ll 1$, разлагая логарифм в ряд, получаем:

$$\Phi_{\text{вент}} = k_0 T \frac{\gamma\beta I}{j_s}, \quad (11-80)$$

т. е. вентильная фото-э. д. с. при малом уровне возбуждения пропорциональна интенсивности света.

Полученные зависимости вентильной фото-э. д. с. от интенсивности возбуждающего света достаточно хорошо согласуются с экспериментальными данными, как это видно из рис. 11-16 и 11-17, на которых приведены вольт-амперная и люкс-амперная характеристики для германиевых фотодиодов в вентильном режиме. На рис. 11-16 видно, что малым токам соответствует предельное для данной освещенности значение фото-э. д. с.

КЛАВДИЯ ВАСИЛЬЕВНА ШАЛИМОВА

ФИЗИКА ПОЛУПРОВОДНИКОВ

Редакторы *В. В. Горбачев, В. А. Абрамов*
Переплет художника *Н. Т. Ярешко*
Технический редактор *О. П. Преснякова*
Корректор *Г. Г. Желтова*

Сдано в набор 21/1 1971 г.
Подписано к печати 5/VIII 1971 г. Т-13343.
Формат 60×90^{1/16}. Бумага типографская № 3.
Усл. печ. л. 19,5. Уч.-изд. л. 20,27.

Тираж 22 000 экз. Цена 94 коп. Зак. № 1567.

Издательство «Энергия». Москва, М-114,
Шлюзовая наб., 10.

Ордена Трудового Красного Знамени Ленинград-
ская типография № 1 «Печатный Двор» имени
А. М. Горького Главполиграфпрома Комитета по
печати при Совете Министров СССР, г. Ленин-
град, Гатчинская ул., 26.